2017

Single-grain OSL chronologies for the Still Bay and Howieson's Poort industries and the transition between them: Further analyses and statistical modelling

Zenobia Jacobs
University of Wollongong, zenobia@uow.edu.au

Richard G. Roberts
University of Wollongong, rgrob@uow.edu.au

Publication Details

Jacobs, Z. & Roberts, R. G. (2017). Single-grain OSL chronologies for the Still Bay and Howieson's Poort industries and the transition between them: Further analyses and statistical modelling. Journal of Human Evolution, 107 1-13.

Single-grain OSL chronologies for the Still Bay and Howieson's Poort industries and the transition between them: Further analyses and statistical modelling

Abstract

The chronology of the Still Bay (SB) and Howieson's Poort (HP) lithic industries remains an issue of keen interest because of the central role of these two phases of technological and behavioural innovation within the Middle Stone Age of southern Africa. Several dating studies have been conducted on SB and HP sites, including a pair published by the present authors and our colleagues in 2008 and 2013. These reported the results of systematically applying single-grain optically stimulated luminescence (OSL) dating procedures to 10 sites in South Africa, Lesotho and Namibia to constrain the timing of the start and end of the SB and HP and reveal the existence of a gap of several millennia between them. Alternative ages for these two industries have since been proposed by others for one of these South African sites (Diepkloof Rockshelter) and some concerns have been raised about the procedures used in our earlier studies to estimate the beta dose rates for a small number of samples. Here, we provide an update on our chronology for the SB and HP and address the issues raised about the methods that we used previously to estimate the beta dose rates and their associated uncertainties. To test the sensitivity of our new SB and HP ages to different underlying assumptions, we have run the same statistical model as that used in our 2008 and 2013 studies under three different scenarios. We show that the ages for the different samples are insensitive to how we analytically process or statistically model our data, and that our earlier conclusions about timing of the start and end of the SB and the HP and the probability of a gap between them remain true for two of the three scenarios. We conclude by bringing our study into the context of additional chronometric, stratigraphic and lithic technology studies that have been conducted in the intervening decade.

Disciplines

Medicine and Health Sciences | Social and Behavioral Sciences

Publication Details

Jacobs, Z. & Roberts, R. G. (2017). Single-grain OSL chronologies for the Still Bay and Howieson's Poort industries and the transition between them: Further analyses and statistical modelling. Journal of Human Evolution, 107 1-13.

Elsevier Editorial System(tm) for Journal of Human Evolution

Manuscript Draft

Manuscript Number: HUMEV-E-16-00501R2

Title: Single-grain OSL chronologies for the Still Bay and Howieson's Poort industries and the transition between them: further analyses and statistical modelling

Article Type: Full Length Article

Keywords: Middle Stone Age; South Africa; optical dating; luminescence dating

Corresponding Author: Professor Zenobia Jacobs, Ph.D

Corresponding Author's Institution: University of Wollongong

First Author: Zenobia Jacobs

Order of Authors: Zenobia Jacobs; Zenobia Jacobs, Ph.D; Richard Roberts, PhD

Abstract: The chronology of the Still Bay (SB) and Howieson's Poort (HP) lithic industries remains an issue of keen interest because of the central role of these two phases of technological and behavioural innovation within the Middle Stone Age of southern Africa. Several dating studies have been conducted on SB and HP sites, including a pair published by the present authors and our colleagues in 2008 and 2013. These reported the results of systematically applying single-grain optically stimulated luminescence (OSL) dating procedures to 10 sites in South Africa, Lesotho and Namibia, to constrain the timing of the start and end of the SB and HP and reveal the existence of a gap of several millennia between them. Alternative ages for these two industries have since been proposed by others for one of these South African sites (Diepkloof Rockshelter) and some concerns have been raised about the procedures used in our earlier studies to estimate the beta dose rates for a small number of samples. Here, we provide an update on our chronology for the SB and HP and address the issues raised about the methods that we used previously to estimate the beta dose rates and their associated uncertainties. To test the sensitivity of our new SB and HP ages to different underlying assumptions, we have run the same statistical model as that used in our 2008 and 2013 studies under three different scenarios. We show that the ages for the different samples are insensitive to how we analytically process or statistically model our data, and that our earlier conclusions about timing of the start and end of the SB and the HP and the probability of a gap between them remain true for two of the three scenarios. We conclude by bringing our study into the context of additional chronometric, stratigraphic and lithic technology studies that have been conducted in the intervening decade.

Response to copy editing

We have looked at and accepted all changes to the document suggested by the copy editor and have specifically done the following changes:

1. provided a corresponding author and e-mail address

2. shortened the abstract to be 313 words

3. swapped on all occasions the Jacobs et al. (2008c) and Jacobs et al. (2008a) citations so that a comes before c

4. changed the Table captions and additions to the footnotes

5. removed the captions from the figures and provide these as Tiff files.

These are two occasions where we chose not to make changes. Both these relate to comments in the main text where the copy editor requested minimal overlap between the text and figure captions. Although there is some overlap, we still feel that there are good reasons for using the wording etc. in the main text to draw out the details of the figures and prefer to leave it as it is.

Single-grain OSL chronologies for the Still Bay and Howieson's Poort industries and the transition between them: further analyses and statistical modelling

Zenobia Jacobs[a*], Richard G. Roberts[a]

[a] ARC Centre of Excellence for Australian Biodiversity and Heritage, Centre for Archaeological Science, School of Earth & Environmental Sciences, University of Wollongong, Wollongong, NSW 2522, Australia.

* Corresponding author: zenobia@uow.edu.au

Abstract

The chronology of the Still Bay (SB) and Howieson's Poort (HP) lithic industries remains an issue of keen interest because of the central role of these two phases of technological and behavioural innovation within the Middle Stone Age of southern Africa. Several dating studies have been conducted on SB and HP sites, including a pair published by the present authors and our colleagues in 2008 and 2013. These reported the results of systematically applying single-grain optically stimulated luminescence (OSL) dating procedures to 10 sites in South Africa, Lesotho and Namibia, to constrain the timing of the start and end of the SB and HP and reveal the existence of a gap of several millennia between them. Alternative ages for these two industries have since been proposed by others for one of these South African sites (Diepkloof Rockshelter) and some concerns have been raised about the procedures used in our earlier studies to estimate the beta dose rates for a small number of samples. Here, we provide an update on our chronology for the SB and HP and address the issues raised about the methods that we used previously to estimate the beta dose rates and their associated uncertainties. To test the sensitivity of our new SB and HP ages to different underlying assumptions, we have run the same statistical model as that used in our 2008 and 2013 studies under three different scenarios. We show that the ages for the different samples are insensitive to how we analytically process or statistically model our data, and that our earlier

conclusions about timing of the start and end of the SB and the HP and the probability of a gap between them remain true for two of the three scenarios. We conclude by bringing our study into the context of additional chronometric, stratigraphic and lithic technology studies that have been conducted in the intervening decade.

Keywords

Middle Stone Age; South Africa; optical dating; luminescence dating

1. Introduction

The Still Bay (SB) and Howieson's Poort (HP) are two widely distributed Middle Stone Age (MSA) industries in southern Africa. Still Bay and HP artefacts have been recovered from sites spread across two million square kilometres of South Africa, Lesotho and Namibia. They exhibit a range of early technological and behavioural innovations, including the complex processing of ochre, the engraving of ochres and ostrich eggshells, the creation of decorative shell beads, and the production of multi-component tools that were hafted using compound adhesives (e.g., Henshilwood et al., 2002, 2004, 2014; Wadley et al., 2009; Texier et al., 2010, 2013; Vanhaeren et al., 2013). Other technological innovations also appear to originate at around the same time (e.g., Brown et al., 2009, 2012; Porraz et al., 2013).

In 2008, we proposed a chronology for the SB and HP, based on a systematic dating study of nine sites spanning a variety of climatic and ecological zones in southern Africa (Jacobs et al., 2008a). Age estimates for 54 sediment samples (and the associated artefacts) were obtained by optically stimulated luminescence (OSL) dating of individual grains of quartz sand, combined with statistical modelling to estimate the start and end dates of the SB and HP. Jacobs et al. (2013) re-examined the timing of the SB, based on the inclusion of 10 samples from Blombos Cave that had been analysed with the same instruments and experimental procedures as those used by Jacobs et al. (2008a).

To discern small differences in age – such as the possible existence of a short time gap between the end of the SB and start of the HP – it is important to estimate ages with maximum precision. Resolving such events and their timing is easily obscured by the chronological 'haze' arising from different dating methods being applied to different sites and, in the case of OSL dating, from the use of different instruments, calibration standards and procedures for sample preparation, measurement and data analysis (Jacobs and Roberts, 2008; Jacobs et al., 2008a). Our published SB/HP chronology is more precise than most other OSL chronologies, because the same instruments and procedures were used to measure all of the samples. This systematic approach

allowed several of the largest uncertainties attached to OSL ages to be removed when comparing our ages against each other, so that a more precise estimate could be made of the time span of the SB and HP industries and any gap between them (Jacobs and Roberts, 2008, 2009; Jacobs et al., 2008c, 2013). That is, the systematic error terms will be the same for each age, so it will not affect comparisons between estimates; only the random measurement errors, listed as the σ_1 values in Fig. 2 of Jacobs et al. (2008a), will contribute to the uncertainty in age. However, when comparing our SB/HP chronology with independent ages – including OSL ages obtained in other laboratories and on other instruments – all of the systematic error terms (listed as the σ_2 values in Fig. 2 of Jacobs et al., 2008a) should be included in the age uncertainty. For readers unfamiliar with the treatment of errors in OSL dating, and in statistics more generally, we refer them to Appendix A in Galbraith and Roberts (2012).

In this paper, we take the opportunity to update our original SB/HP data set using recent updates in dose rate conversion factors and error estimation of beta dose rates (Jacobs and Roberts, 2015), and to then re-run the age model for further comparison. In doing so, we confirm the high likelihood of a gap of several millennia between the SB and HP for those samples reported by us in 2008, and demonstrate that the published single-grain OSL chronologies for these two industries (Jacobs et al., 2008a, 2013) are robust relative to a range of alternative assumptions and modelled ages. We also discuss the two aspects of our original analysis that were commented on by Guérin et al. (2013): (1) the small size of the uncertainties attached to the beta-particle contribution to the environmental dose rate; and (2) the details of implementation of a method proposed by Jacobs et al. (2008b) to adjust the beta dose rate in specific circumstances. The latter adjustment affected a subset of the samples from three of the nine sites reported in Jacobs et al. (2008a). We briefly discuss the first issue below (see section 4), but have addressed it in full in Jacobs and Roberts (2015), together with a new method that is now used in our laboratory to calculate beta dose rate errors when using our GM-25-5 beta counters. Here, we focus on the concept and implementation of the beta dose rate adjustment method of Jacobs et al. (2008b) to explore the effect of different

approaches – including that advocated by Guérin et al. (2013) – on the ages calculated for the SB and HP samples. We also discuss the previously published data set from Pinnacle Point site 5-6 (PP5-6) that contains the HP, as well as evidence of a superficially similar microlithic industry earlier in time (Brown et al., 2012). The chronology of the latter is starting to fill the 'gap' between the SB and the HP, and, hence, directly address the question of the timing of the transition between the SB and the HP. This and other studies on lithic technology and chronology in this critical time period are currently stimulating a reassessment of the meaning of the HP (e.g., Mackay et al., 2011a,b, 2014; Porraz et al., 2013; Conard and Porraz, 2015) and the SB (e.g., Archer et al., 2015, 2016) as homogeneous entities at different time scales.

2. Background

2.1 OSL dating and field context

There are two parts to the OSL dating equation; Jacobs and Roberts (2007), Duller (2008), Wintle (2014) and Roberts et al. (2015) provide overviews of OSL dating for non-specialists. The numerator is called the equivalent dose (usually denoted as D_e and expressed in gray, Gy), which represents the radiation energy absorbed by a mineral grain since it was last exposed to sunlight ('bleached') or heated to a high temperature. The D_e is estimated from measurements of the OSL signal emitted by individual grains or by single aliquots composed of multiple grains. The burial history of an individual grain is the smallest meaningful unit of analysis in OSL dating, so Jacobs et al. (2008c) obtained D_e estimates for all of the samples from OSL measurements of individual sand-sized grains of quartz.

The denominator in the OSL age equation is the environmental dose rate (Gy per unit time), which consists of four separate contributions: the beta-particle and gamma-ray dose rates from materials surrounding the quartz grains by up to ~3 mm and ~30 cm, respectively, and lesser contributions from cosmic rays and alpha particles (the latter emitted by radioactive inclusions internal to the quartz grains). Many archaeological deposits are heterogeneous in their composition

and in the spatial distribution of the organic and inorganic constituents (e.g., Goldberg et al., 2009; Miller et al., 2013; Karkanas et al., 2015). Changes also commonly occur over time due to a variety of diagenetic, pedogenic and taphonomic processes. As a result, quartz grains deposited at the same time and situated within a few mm or cm of each other can have very different beta dose rates, depending on their relative proximity to materials of high or low radioactivity. Quartz itself typically has a very low dose rate, as do many secondary carbonates formed in archaeological sediments after deposition. By contrast, minerals such as zircon and potassium feldspar have much higher dose rates, so a quartz grain juxtaposed by a zircon grain will receive a much higher beta dose rate than a contemporaneous quartz grain coated in calcium carbonate, for example.

The differences in beta dose rate to individual quartz grains is difficult to model in a way that is both simple to implement and faithful to the field context of the sample, because each sample poses a unique set of circumstances that no single model can address without knowing the boundary conditions specific to that sample. Ideally, what is required is knowledge of the beta dose rate to each dated grain in its original, undisturbed burial position. Combining spatially resolved measurements of the OSL signal from individual grains with measurements of the dose rate at the single-grain scale of analysis is currently a research priority (e.g., Martin et al., 2015a,b; Roberts et al., 2015). At the present time, however, it is not feasible to model the beta dose rate to individual grains for each and every sample.

In practice, sediment samples are disaggregated in the laboratory and quartz grains are separated for OSL measurements. The dose rate is determined for the bulk sample (i.e., the sample as collected from the field) and D_e values are obtained from single grains or multi-grain aliquots composed of purified quartz grains. Based on the extent of scatter among these D_e values, and the existence of any patterns in the D_e distribution, some estimate is made of the overall D_e value that corresponds most closely to the last bleaching or heating event. Several well-established statistical models are widely used in OSL dating for combining independent D_e values appropriately, with the choice of model depending on sample context, among other things (Galbraith and Roberts, 2012).

The D_e so obtained is representative of the values for the group of grains in the sample that are thought to be related to the event of interest, and this population-estimate of D_e is then usually divided by the bulk sample ('sample-average') dose rate to determine the OSL age. The latter is, thus, not a single-grain age in the strict sense, because the sample-average dose rate is used for all grains in the population, and not the dose rate specific to each grain.

For samples consisting of grains that were well bleached before burial and that have not been affected by post-depositional mixing, the central age model (CAM) is often used to determine the weighted mean D_e for all of the grains (Galbraith et al., 1999); the CAM D_e is divided by the sample-average dose rate to estimate the OSL age. When calculating the standard error on the D_e, the CAM includes an allowance for the spread in D_e values over and above their measurement errors. This additional 'overdispersion' (OD) is a ubiquitous feature of D_e distributions. For quartz grains that had been fully bleached at the time of deposition and thereafter remained undisturbed, the single-grain D_e distributions are typically overdispersed by 10–20% or more (e.g., Galbraith et al., 2005; Arnold and Roberts, 2009).

The CAM may be sufficient for mineral grains buried in homogeneous radiation environments, because all of the grains will have experienced similar dose rates during the period of burial. It may even be appropriate for samples in which the dose rates to individual grains are markedly different, provided the spread in single-grain D_e values is the same as the distribution of grain-specific dose rates. Suffice to say here that the CAM and sample-average dose rate can provide reliable estimates of age for many samples – indeed, Jacobs et al. (2008a) used this combination for many samples in their SB/HP study. But it is not the only approach that can be used, nor necessarily the best, and this is especially true for grains buried in heterogeneous radiation environments (Galbraith, 2015b).

2.2 Beta dose rate heterogeneity

At many archaeological sites, such as Sibudu Cave, the radiation environment is far from uniform at the scale of beta particles, because of spatial differences in the texture and composition of the deposit over distances of a few millimetres. Microscopic and spectroscopic studies illustrate the complex micromorphology of the Sibudu deposits, including the existence of a variety of microstratigraphic structures associated with burnt material and bedding (e.g., Schiegl et al., 2004; Schiegl and Conard, 2006; Goldberg et al., 2009; Wadley et al., 2011). Most of the deposit is anthropogenic in character, with lesser amounts of geogenic and biogenic material (Goldberg et al., 2009). Given the heterogeneity of the deposit at the scale of an individual sand grain, it is not unreasonable to expect that quartz grains deposited at the same time will experience different beta dose rates after burial and that these grain-to-grain differences will contribute to the spread in single-grain D_e values.

Jacobs et al. (2008c) measured the D_e values for individual grains of quartz extracted from 14 sediment samples collected from the post-HP, late MSA and final MSA deposits at Sibudu and reported three types of D_e distribution: one consistent with a single dose population, another indicative of mixing between the MSA and overlying Iron Age deposits, and a third that they termed 'scattered'. Most of the samples had scattered D_e distributions, which were distinguished by a small population of grains with much smaller D_e values than the majority (>82%) of grains in the sample. Jacobs et al. (2008c) attributed the minor component to those grains experiencing smaller beta dose rates than the majority, as would be expected if the grains had been surrounded by >2 mm of radioactively inert material, which occur in the deposits (Schiegl et al., 2004; Schiegl and Conard, 2006; Goldberg et al., 2009). Jacobs et al. (2008b) assumed that dividing the D_e values for these grains (i.e., the minor component of grains with smaller D_e values) by the sample-average dose rate would give rise to age underestimates, and thus devised a procedure to correct for this bias. We note that such D_e distributions and reasons for spread could be interpreted differently and may well change if more data are collected and further archaeological and geological studies are conducted at a site.

The 3-step procedure of Jacobs et al. (2008b) first consisted of fitting the finite mixture model (FMM) to the D_e distributions to estimate the D_e value of each population of grains and the relative proportion of grains in each fitted component (Roberts et al., 2000; Galbraith and Roberts, 2012). Then, the beta dose rate of the minor component was calculated by making a simple assumption: that it cannot be less than zero or more than the sample-average beta dose rate. A value mid-way between these two extremes was chosen as a first approximation, and assigned a standard deviation that spanned this full range at 2σ. The D_e component that contains the majority of grains represents the population of interest, so the final step of the procedure was to estimate the beta dose rate for this major component. This was accomplished by making a pro rata adjustment to the sample-average beta dose rate based on the proportion of grains in the major and minor components. So, the sum of the beta dose rates for both components (taking into account the proportion of grains in each) equals the sample-average beta dose rate. For these 13 samples, the ages were thus obtained by dividing the D_e of the major component by the dose rate specific to that population of grains. (See Jacobs et al. [2008b: section 10.4] and Jacobs et al. [2011, their Table 4] for details of this approach and worked examples.)

At Sibudu, Jacobs et al. (2008b) proposed two reasons, field context and statistical precision, to use the FMM D_e values and adjusted beta dose rates (henceforth 'aFMM ages') to estimate the OSL ages for these samples, rather than the CAM D_e values and sample-average beta dose rates (henceforth 'CAM ages'). The CAM ages calculated by Jacobs et al. (2008b) for all of their samples produced a conspicuous outlier among the post-HP samples: the age of SIB9 was much too young for its stratigraphic position, and the finely-resolved stratigraphy ruled out mixing as the cause. The 'homogeneity test' (Galbraith, 2003; Galbraith and Roberts, 2012) suggests that the CAM ages (and associated random errors) of all six post-HP Sibudu samples do not come from a single population ($p = 0.0045$). This appears to be due to the inclusion of the SIB9 age, which differs by more than 3σ ($p = 0.0006$) from the weighted mean CAM age for the other five post-HP samples. Sample SIB9 was collected from stratum P, "an orange–brown sandy-silt with white flecks

of gypsum" and small ashy lenses that probably represent the remains of a hearth (Schiegl and Conard, 2006, their Table 1). Given the observations of gypsum in this and other strata at Sibudu, there were prima facie contextual grounds for expecting that some quartz grains in SIB9 – and, by extension, other Sibudu samples with similarly scattered D_e distributions – would have received much lower beta dose rates than the majority. This, of course, is a very simplified assumption and if we scrutinise the distribution of radioactivity on a microscale more closely we will doubtless see further variation due to a range of different factors. But given the constraints of their knowledge at the time, Jacobs et al. (2008b) applied their dose rate adjustment procedure to these samples and obtained aFMM ages without any significant stratigraphic anomalies. Accordingly, Jacobs et al. (2008b,c) proposed that the aFMM ages were more accurate than the CAM ages. Three of the 14 aFMM ages were also more precise than their CAM counterparts. This outcome can occur because the D_e may be more precisely estimated using the FMM than the CAM, and this increase in precision may outweigh the decrease in precision associated with the adjusted beta dose rate. The balance between these different factors will differ between samples; see Galbraith (2015a,b) for related statistical discussion.

Except for SIB9, there are negligible differences between the aFMM and CAM ages for the Sibudu samples dated by Jacobs et al. (2008c). Figure 1a shows the CAM and aFMM ages measured for 13 post-HP, late and final MSA samples from Sibudu (the age of SIB12 was determined using the CAM only). Four systematic errors are not included for comparison, because they are the same for both (aFMM and CAM) data sets. These are the errors associated with the laboratory reference standard and the grain-size attenuation factors used to estimate the beta dose rates (see Appendix A in Jacobs and Roberts, 2015), the systematic error on the cosmic-ray dose rates, and the allowance of 2% for possible bias in the calibration of the laboratory beta source used for D_e determination.

The same data are shown in Figure 1b as the difference between each pair of log age estimates plotted against the mean log age of the pair (after Oldham, 1962). We have used the

natural logarithms of the ages, following Galbraith and Roberts (2012), because the size of the standard errors increases in proportion to the ages. It is difficult to calculate the standard error of a difference in log age estimates from the individual standard errors in Fig. 1a because the aFMM and CAM ages are both obtained from the same data, so their estimation errors are correlated, but to an unknown degree. However, the standard deviation of the differences gives an estimate of this individual standard error, assuming it is approximately the same for each pair of ages. The grey band in Figure 1b shows 2.18 of these standard deviations above and below zero, where 2.18 is the 97.5 percentage point of the t-distribution with 12 degrees of freedom (equivalent to 1.96 for a normal distribution), so that any point within this band has a 95% confidence interval for the true difference that includes zero. The ratios of the aFMM/CAM ages are consistent with unity for all samples, except for SIB9 which lies away from the 1:1 line (Fig. 1a) and outside the grey band (Fig. 1b).

No model should be applied 'blind' to any sample, without contextual support for its use (Galbraith et al., 1999, 2005; Galbraith and Roberts, 2012). It is for this reason that we carry out intensive fieldwork alongside archaeologists, geoarchaeologists and, where possible, with those doing micromorphology and related studies, to understand better not only the site formation processes at play at each site, but also the finer scale differences between the sedimentary and archaeological characteristics of different units and layers. In the absence of a specific reason to use the FMM or some other model (such as the minimum or maximum age models; Galbraith and Roberts, 2012), and taking into account knowledge of site and sample context, the CAM D_e divided by the sample-average dose rate may give the most reliable estimate of age. Such was the case with some or all of the OSL samples from Apollo 11, Klasies River, Melikane, Ntloana Tsoana, Rose Cottage Cave, Sehonghong and Sibudu dated by Jacobs et al. (2008a), and for seven (of 10) samples associated with the SB at Blombos (Jacobs et al., 2013). For samples with mixed and/or scattered D_e distributions, there is nonetheless merit in also exploring other options (Galbraith, 2015a,b).

3. Methods

3.1 Dose rate determination

In this study, we have recalculated the dose rates for the SB and HP samples, incorporating: (1) the updates to dose rate conversion factors reported in 2011 (Guérin et al., 2011); (2) the error estimation procedure for beta dose rates published by Jacobs and Roberts (2015); and (3) the uncertainty associated with the water content of each sample, to which we assigned a minimum relative standard error of ± 25%. These changes affect our estimates of the beta dose rate and, in some cases, also the gamma dose rate (with, in general, a slight increase in dose rate and, hence, decrease in age). For some of the samples, we also re-measured the beta dose rates using the GM-25-5 beta counters and used the average of the new and old measurements, where consistent. We found an inconsistent result for only one of the samples (DRS6) and used the new results (which we confirmed by replication) to calculate the revised age reported in Jacobs and Roberts (2015).

Our revised method of error estimation for beta dose rates was formulated with statistician Rex Galbraith and is described in full in Appendix A of Jacobs and Roberts (2015), together with a worked example for one of the samples from Diepkloof Rockshelter. The procedure is based on an analysis of variance between replicate measurements that is applicable to all samples measured using our GM-25-5 beta counters; we have always counted at least three replicates of each sample to check for reproducibility. Our error estimation procedure takes into account the counting errors associated with day-to-day fluctuations in the behaviour of the equipment, the differences in counting efficiencies of the five detectors, the reproducibility among the three replicates measured for each sample, and the counting statistics of the replicates, the laboratory standard and the background blank. The total error on the beta dose rate also includes systematic errors for the measurement imprecision of the laboratory reference standard (~1.8%), the uncertainties associated with the dose rate conversion factors (assumed to be ~2%) and a ~2% uncertainty on the correction required for grain-size attenuation of beta particles (Murray and Olley, 2002). There may well be

other sources of uncertainty that can be included or their estimation improved. The purpose of publishing our method is to show exactly what we include and how we calculate the errors associated with our beta dose rates. We encourage others to do likewise for transparency and to enable further improvements to be made to error estimation.

3.2 Chronological modelling

We have fitted the same statistical model as that used by Jacobs et al. (2008a, 2013) to the revised ages to estimate new start and end ages for the SB and HP, the duration of the HP period, and the length of the gap between the SB and HP. We tested three different scenarios to examine the sensitivity of the modelled ages to a range of alternative assumptions:

Estimate A uses all the updated ages for all of the samples, with the aFMM ages used instead of the CAM ages for those samples where the scatter in D_e values is thought to be due to dose rate heterogeneity; these are indicated with asterisks in Supplementary Online Material (SOM) Table S1. For sample DRS6, the FMM age is used in preference to the CAM age (see Jacobs and Roberts, 2015 for discussion of this sample), but the model outcome is not sensitive to this choice. The individual ages are listed in SOM Table S1. This estimate represents the most direct comparison between the revised ages and the ages reported by Jacobs et al. (2008a, 2013).

Estimate B uses the same samples as A, but CAM ages with sample-average beta dose rates are used for the 10 samples (three from Blombos and seven from Sibudu) for which aFMM ages are used in A. The individual ages are listed in Table 1 and SOM Table S1; these now represent our preferred age estimates. This estimate is also the analysis closest to that favoured by Guérin et al. (2013).

Estimate C is the same as B, except that three of the Diepkloof samples (DRS11, 13 and 14) are omitted due to the currently inexplicable differences in age for the middle part of this MSA sequence (Tribolo et al., 2013; Feathers, 2015; Jacobs and Roberts, 2015).

As in Jacobs et al. (2008a, 2013), we calculated standard errors σ1 and σ2 (Table 1 and SOM Table S1). These exclude and include, respectively, systematic errors that are common to each of our age estimates and, hence, do not affect comparisons among estimates. The σ1 values exclude the four systematic errors noted earlier (i.e., those associated with the laboratory beta source and GM-25-5 beta counter reference standard, the grain-size attenuation factors and the cosmic-ray dose rates). As the gamma dose rates were measured using a variety of methods at the different sites, the associated uncertainties are included in σ1. Uncertainties in water content are also treated as random errors and, thus, included in σ1.

4. Results and preliminary discussion

In this section, we present the results of these new analyses, starting with an assessment of the accuracy and precision of the beta dose rates estimated using our GM-25-5 beta counters and the procedures described in Jacobs and Roberts (2015). This is followed by a discussion of the merits of this approach in the light of the remarks made by Guérin et al. (2013). We then present the results of applying the statistical model to Estimates A, B and C to ascertain the chronology of the SB and HP industries, and discuss the strength of the evidence for a time gap between them.

4.1 Beta-counting accuracy and precision

The accuracy of our GM-25-5 beta counters and sample preparation and presentation procedures has been tested and confirmed in several ways, using the samples from Diepkloof (see Jacobs and Roberts, 2015). We have measured samples as both loose powder and as powder pressed into pellets that are bound together with a few drops of polyvinyl acetate (PVA), with and without delay times between preparation and measurement (to allow for equilibration of the ^{238}U and ^{232}Th decay chains). For the 13 samples from Diepkloof, the beta dose rates obtained for the loose sediment (no delay time) are, on average, ~2.4% higher than the corresponding results for the pressed pellets (with a one month delay time; ratio of 1.024 ± 0.017), but the results are randomly

spread around a ratio of unity and are not systematically biased (Jacobs and Roberts, 2015: Fig. 8a). We have also measured the U, Th and K contents of the same samples using other techniques and obtained consistent results. For the Diepkloof samples, we compared the beta dose rates obtained for the pressed pellets (with a one month delay time) using our GM-25-5 beta counters with measurements of U, Th and K from: (1) inductively-coupled plasma mass spectrometry (ICP-MS) and optical emission spectroscopy (ICP-OES); and (2) thick-source alpha counting and X-ray fluorescence (XRF) spectroscopy. Average ratios consistent with unity (0.996 ± 0.015 and 0.997 ± 0.015, respectively) were obtained for these sets of data, again with individual ratios spread randomly around the 1:1 line (Jacobs and Roberts, 2015, their Fig. 8b,c). The ICP-MS, ICP-OES and XRF measurements were made in commercial laboratories, so they provide independent validation of the accuracy of our GM-25-5 beta-counting measurements.

Jacobs et al. (2008a,b,c, 2011, 2012, 2013) also used GM-25-5 beta counting to determine the beta dose rates for the majority of their samples. Guérin et al. (2013) considered the reported errors to be too small and instead suggested that a beta dose rate error of ~7% may be more appropriate, based on their analysis of measurements made by Ankjærgaard and Murray (2007). We note that the beta dose rates reported in six of the seven papers referred to in SOM Table S1 of Guérin et al. (2013) actually have errors that are similar to a value of 7%: 5.2–7.4% (El Harhourha 2, Morocco), 5.6–8.3% (El Mnasra, Morocco), 5.1–7.0% (Contrebandiers, Morocco), 5.1–6.6% (Mumba, Tanzania), 5.8–7.3% (Blombos, South Africa), 5.2–8.5% (Pinnacle Point site 13B, South Africa) and 2.6–9.9% for the post-HP, late and final MSA samples from Sibudu, South Africa. Guérin et al. (2013) do not acknowledge this fact, but imply instead that the beta dose rate errors associated with all of the ages are too small. Only in Jacobs et al. (2008a) are the beta dose rate errors for the SB and HP samples much smaller than 7%, ranging between 1.1% and 3.8%, and we concur with Guérin et al. (2013) that these errors are underestimated. We re-examined the beta dose rate errors published in Jacobs et al. (2008a) and found that we had not fully propagated all errors for these samples. This oversight is rectified in the current study. Our revised estimates of the total

error (random plus systematic) on the beta dose rates range from 4.9 to 7.8% for the SB and HP samples in this study; the random errors alone range between 2.0 and 6.8%, which are larger than those reported by Jacobs et al. (2008a). Since we have undertaken an evaluation of the specific sources of uncertainties in our beta counting measurements (see Jacobs and Roberts, 2015), and because conclusions obtained from experiments using a particular set of settings or circumstances are strictly valid only for those specific conditions (Galbraith and Roberts, 2012), we use our revised estimates in preference to the 7% value suggested by Guérin et al. (2013).

4.2 Updating the SB–HP chronology

The updated dose rate data and ages are presented in Table 1; these ages supersede all previously reported ages for these samples. The modelled estimates for the start and end ages of the SB and HP are listed in Table 2, while those for the length of the HP period and the gap between the SB and HP are given in Table 3; the original results of Jacobs et al. (2008a) are also provided in both tables, for comparison. The confidence intervals for the start and end ages of the SB and HP include all random and systematic errors (i.e., $\sigma 2$), so these ages can be compared directly with other independent chronologies. To estimate the confidence intervals for the durations in Table 3, all systematic errors have been omitted (i.e., $\sigma 1$).

4.2.1 Start and end ages for the SB and HP and the length of the HP Start and end ages for both the SB and HP are very similar for Estimates A, B and C (Table 2) and agree closely with the original estimates (Jacobs et al., 2008a). The length of the HP has increased under all three scenarios, from the original estimate of 5.3 ka to 7.7 ka (Estimate A), 8.4 ka (Estimate B) and 8.6 ka (Estimate C) (Table 3). Based on our preferred ages (Estimate B), the insensitivity of the modelled estimates to these different scenarios, and the numerous updates and new measurements we have made, we remain confident in our age estimates for the HP and SB (Tables 1 and 2). We also note the good agreement between our ages for the SB and HP and those obtained using other methods and by

other laboratories (e.g., Miller et al., 1999; Grün et al., 2001; Vogel, 2001; Tribolo et al., 2005, 2006; Valladas et al., 2005; Cochrane, 2009; Högberg and Larsson, 2011; Henshilwood et al., 2014; Murray-Wallace et al., 2015).

We caution against conflating these small changes in estimated ages for the SB and HP with issues associated with the chronology for Diepkloof. In Estimate C, we deliberately excluded three samples collected from the middle part of the Diepkloof sequence, because of the unaccountable differences currently observed in three separate studies. We note, however, that our age estimates for the SB and HP at Diepkloof are consistent with those obtained at other sites, and that their exclusion from the model has little effect on the calculated start and end ages for the HP and SB or on the estimated length of the HP. We show below (see crosses in Fig. 2) that the chronological differences in this part of the Diepkloof sequence are not due to the application of the beta dose rate adjustment method (as was suggested by Guérin et al., 2013), but are likely related to the measurement of potassium (^{40}K) in these deposits (Jacobs and Roberts, 2015). Other aspects of the discrepant chronology for the middle part of the Diepkloof sequence have previously been addressed by Jacobs and Roberts (2015).

4.2.2 The gap between the SB and the HP Jacobs et al. (2008a) identified a gap between the SB and HP (p = 0.008) and estimated its length to be 6.7 ka, with a 95% confidence interval of 2.7–9.3 ka (Table 3). The gap durations for Estimates A, B and C are also shown in Table 3, together with the 95% confidence intervals and p-values. The latter are based on the likelihood ratio test of the null hypothesis that there is no gap between the end of the SB and the start of the HP. The length of the gap for Estimates A, B and C are all slightly shorter than estimated originally. The 95% confidence interval for Estimate A includes zero, so those data are consistent (at the conventional 5% significance level) with the null hypothesis that there is no gap; and also, of course, with the hypothesis that there is a gap of up to 8.3 ka. The p-value of 0.059 (i.e., slightly greater than 0.05) suggests only weak evidence in favour of a gap. Estimates B and C show clear statistical evidence

in favour of a gap of ~5 ka, formally supported by *p*-values of 0.014 and 0.033, respectively. As noted above, our preferred ages are used in Estimate B, which is also the scenario closest to that favoured by Guérin et al. (2013).

The general conclusions of Jacobs et al. (2008a), thus, remain true under two of the three scenarios (Estimates B and C), both with regards to the general timing of the start and end of the SB and the HP, and the likely gap between them of several millennia. Our results are not an artefact of how we analytically process or statistically model our data.

5. Discussion

In this section, we discuss the main concerns raised by Guérin et al. (2013) about the beta dose rate adjustment procedure used for a subset of the samples dated by Jacobs et al. (2008a). We have demonstrated above that the modelled chronology for the SB and HP is not sensitive to the inclusion or exclusion of the ages obtained using this procedure, but it is important nonetheless to address some of the statements made by Guérin et al. (2013). We conclude by reassessing the existence of a gap between the SB and HP in view of archaeological developments that have taken place since our original study (Jacobs et al., 2008a), and we highlight the importance of temporal resolution (scales of analysis) when interpreting OSL (and other) chronologies in the context of stratigraphic and archaeological changes.

We note that almost all of the examples in Guérin et al. (2013), including the only figure, are unrelated to the timing of the SB and HP or of the transition between them – and, hence, to the title of their commentary. Instead, they concentrate on the post-HP, late and final MSA at Sibudu and on four sites in East and North Africa: Mumba Rockshelter in Tanzania (Gliganic et al., 2012a,b) and three caves in Morocco (La Grotte des Contrebandiers, El Harhoura 2 and El Mnasra; Jacobs et al., 2011, 2012). Interested readers are encouraged to read our original papers, and the supplementary information sections, for detailed accounts of the procedures used and the data so obtained in these studies.

5.1 Beta dose rate adjustments

5.1.1 "All models are wrong but some are useful" This is a famous quote by the statistician George Box (Box, 1979: 202) and reflects a truism: models are a simplification of the real world to a greater or lesser extent, but some models can provide useful approximations and illuminating insights. In the context of OSL dating, a dose rate model is only useful if it bears some (close) relation to the depositional environment of the samples of interest and if the sample was collected and measured in an appropriate manner.

The beta dose rate adjustment method of Jacobs et al. (2008b) was predicated on two observations: (1) the scattered D_e distributions of some of the Sibudu samples, which could not be explained by incomplete bleaching of the grains before burial or mixing after burial; and (2) the presence of low-radioactivity minerals (gypsum and calcite) in the deposit. This model involved some simplifying assumptions, such as estimating the number of single-grain populations in any particular D_e distribution to choose between applications of the CAM or (a)FMM. On the basis of two equations, Guérin et al. (2013: 315) claimed that the model proposed by Jacobs et al. (2008b) was "flawed and unnecessary" and would "introduce additional uncertainties", concluding that the average D_e and dose rate would give the best age estimate for well-bleached samples with beta dose rate heterogeneity.

Galbraith (2015a) explained in detail how the assumption underpinning the two equations of Guérin et al. (2013) is not tenable and not true, in general, for measured quantities such as D_e estimates, and showed that their conclusion does not follow from their two equations – whether or not they hold in practice. We agree with Galbraith (2015a) that the initial assumption underlying the equations of Guérin et al. (2013) – that the only source of dispersion in single-grain doses is the dose rate – was incorrectly attributed to Jacobs et al. (2008b). On the contrary, Jacobs et al. (2008b) listed numerous factors that contribute to the dispersion in single-grain D_e values, including intrinsic factors related to the OSL characteristics and behaviour of individual grains, and extrinsic

factors related to the depositional and post-depositional histories of the sample. They referred to differences in the beta dose received by individual grains in their burial environment and the inadequate bleaching of some grains before burial, as well as the potential for post-depositional movement of grains, and also stated that some dispersion in single-grain D_e values may arise from non-identical field and laboratory conditions and natural variability (following Galbraith et al., 2005).

5.1.2 Age model selection Some of the other statements of Guérin et al. (2013: 314) are also unfounded. They claimed that Jacobs et al. (2008a) choose between the CAM and FMM on the basis of a "potentially subjective 20% limit" for D_e OD. We have never used an OD value of 20% as the sole criterion for choosing between the CAM and any other age model, but we acknowledge that our wording may have been ambiguous if read in isolation from any of our other papers. Instead, we routinely take several steps to assess the possible cause(s) of the observed distribution of D_e values, as we have described in previous publications (e.g., Jacobs and Roberts, 2007, 2015; Jacobs et al., 2008c, 2011, 2013). First, we calculate the OD value for each sample using the CAM. Overdispersion values of up to 20% or more have been frequently observed for many quartz samples from around the world (e.g., Olley et al., 2004; Arnold and Roberts, 2009). So, when OD values of ~20% or greater are obtained, we give special consideration to any other features of the D_e distribution, but do not automatically apply the FMM or minimum age model (MAM). Second, we make a visual appraisal of the distribution pattern of D_e values in a radial plot (Galbraith, 1988; Galbraith and Roberts, 2012) to assess whether it comprises a single population of values or multiple, discrete components, or if it displays some other feature(s) of interest. Third, we use our knowledge of the site and sample contexts to ask questions about the known depositional and post-depositional processes that could give rise to the observed D_e distribution. These different threads of evidence are then brought together to develop a logical scenario to explain the D_e distribution for any particular sample and to help us select the most appropriate statistical model for age estimation. We note that different practitioners may interpret the same datasets differently. No interpretation is

flawless, but must be explained so that the readers of a paper can make up their own minds and to enable data sets to be re-examined in the light of new observations and associated archaeological and geological information.

When the FMM is selected, we always use statistical diagnostics, namely the Bayes Information Criterion (BIC) and the maximum log likelihood (llik), based on the recommendations of Galbraith (2005). This permits us to decide: (1) whether a D_e distribution is made up of more than one discrete D_e component; and (2) if more than one discrete component can be statistically supported, then the BIC and llik are used to resolve, for a range of tested alternatives, the number of D_e components and OD value at which the model is optimised. Jacobs et al. (2008c) reported OD values of 7% for grains that were bleached and then given a uniform beta dose in the laboratory (for a dose recovery test on 1000 single grains) and 17% for grains from a natural sample (SIB12) that consisted of a single D_e component. These represent 'best case' outcomes for laboratory- and field-dosed samples composed of well-bleached grains. Jacobs et al. (2008c) fitted the FMM to the D_e distributions of the other Sibudu samples using a conservative range of OD values (10–20%), and noted that very similar D_e estimates were obtained for all OD values within this range. We did not fit the FMM with OD values of >20% because optimised values of llik and BIC were obtained for the model at values within the 10–20% range. The OD value at which the model was optimised was reported in Jacobs et al. (2008b,c), alongside a worked example, and in Jacobs et al. (2008a) for each sample where the FMM was used. A wider range of values may well be necessary for samples from other sites, but this needs to be assessed on a case-by-case basis. Therefore, contrary to the above claim of Guérin et al. (2013), we maintain that our implementation of the FMM is statistically robust and suitable for our samples.

Guérin et al. (2013: 314) also remarked that the intrusive grains in the three mixed Iron Age/MSA samples were discarded "without explicit justification". This is not true. Jacobs et al. (2008b,c) stated that those grains were discarded because their D_e values were too small to be explained by low beta dose rates and yielded ages consistent with the Iron Age. Jacobs et al.

(2008b) also provided additional evidence of post-depositional mixing for these three samples, based on the extent of D_e overdispersion among aliquots composed of ~50 grains. Clearly, it would be nonsensical to retain Iron Age grains in data sets used to determine the burial ages of MSA deposits.

5.1.3 Dose rate simulations The beta dose rate adjustment method of Jacobs et al. (2008b) was also challenged by Guérin et al. (2013) on the basis that bimodal D_e distributions were not generated in the numerical (Monte Carlo) simulations of Nathan et al. (2003). However, the imposed constraints of these simulations make them poor approximations of many real archaeological deposits – in particular, the assumed random distribution of radioactively inert or comparatively high radioactivity spheres in an otherwise homogeneous matrix. No allowance was made for spatial heterogeneity in beta dose rate associated with common features such as hearths, organic matter and the precipitation of secondary minerals around grains and in the voids between them. Non-randomly distributed features at the microscopic scale, such as laminated layers, have been comprehensively documented at Sibudu (Goldberg et al., 2009; Wadley et al., 2011), but they were not incorporated in the simulations. Indeed, Nathan et al. (2003: 312) recognised that while they were unable to simulate bimodal distributions, grains surrounded completely by carbonate material "may be an exception to this." Likewise, Cunningham et al. (2012) conducted an experimental and Monte Carlo simulation of beta dose rate heterogeneity among laboratory-dosed quartz grains, and found that the simulation underestimated the spread in measured single-grain D_e values. Various possible causes of the discrepancy were considered, including those that "lie in the parts of the real world that are not included in the model" (Cunningham et al., 2012: 1067).

Our opinion is that generalised models of beta microdosimetry are not likely to prove useful (sensu Box, 1979) in the absence of specific information about the burial environment. That is, there is no a priori basis to suppose that they will provide a realistic approximation of the beta dose rate for any particular sample. A 'useful model' requires knowledge of the field context of the

samples under investigation to ensure that any spatial variability in the beta dose rate is taken into account – that is, the model must allow for all things that are known to be true, rather than be developed in isolation from physical reality. There is merit in measuring the spatial distribution of radioactivity in the samples of interest and then constructing a numerical model based on those data. The beta dose rate adjustment method of Jacobs et al. (2008b) attempted to incorporate knowledge of field context for samples that are known, or are thought, to contain a proportion of low dose rate material. More sophisticated models, coupled with Monte Carlo simulations (e.g., Martin et al., 2015a,b), will offer advantages if the field context supports their application and if the necessary data are available to implement them in practice. Accomplishing the latter will be non-trivial, however, if the burial environment is complex (see Roberts et al., 2015).

The fact that the Sibudu samples with scattered D_e distributions could be optimally fitted by only two FMM components is, of course, a simplification of reality. For all of the non-single D_e distributions, Jacobs et al. (2008c: 1800) noted the common occurrence of "a small number of grains with intermediate values", but the vast majority of grains in the scattered D_e distributions were captured by just two components, thus representing a useful approximation. Monte Carlo simulations are also an approximation of reality and all relevant parameters have not yet been included in simulations of natural samples. This shortcoming should be rectified in future simulations, so that models are able to explain the diversity of D_e distributions observed in real samples. Likewise, interpretation of complex D_e distributions should ideally be accompanied by supporting micromorphological observations and chemical analyses of the related sediments (e.g., Jankowski et al., 2015).

5.1.4 Age comparisons Guérin et al. (2013) contend that the aFMM ages of Jacobs et al. (2008b) are less accurate than the CAM ages and, on p. 315, they say that the CAM and aFMM ages are significantly different "for 13 out of 14 samples at Sibudu." In fact, the numbers are the reverse: only one sample (SIB9) out of the 14 has an aFMM age that differs significantly from its CAM age

(see Fig. 1a,b). Of course, the CAM age of SIB9 identified as an outlier by Jacobs et al. (2008b) provided the catalyst for their beta dose rate adjustment method. None of these samples is from the HP or SB, so any differences between the two methods will have very little effect on the calculated chronology provided in Jacobs et al. (2008a). The same applies to the North African data set of Jacobs et al. (2012) that Guérin et al. (2013) also examined. Thirty-two OSL samples were dated from two sites in Morocco (El Harhoura 2 and El Mnasra) and 20 of these samples had scattered D_e distributions. As noted earlier, it is difficult to calculate a valid standard error for the difference or ratio of the aFMM and CAM ages from their individual standard errors, because of the unknown correlation between their estimation errors. A simple valid comparison can be made using a standard paired t-test and confidence interval applied to the logs of the ratios. For these 20 samples, the mean of the log ratios is 0.021, corresponding to an aFMM/CAM ratio of 1.022 (95% confidence interval from 1.008 to 1.035). So, on average, the aFMM ages are about 2% higher than the CAM ages. The individual log ratios vary between –0.052 and +0.064 (equivalent to ratios of between 0.949 and 1.066), with a standard deviation of 0.028. The latter figure can be interpreted as an approximate average relative standard error of 2.8% applicable to a single aFMM/CAM age ratio. These differences are small, given that the relative standard errors of the aFMM and CAM ages are typically ~5% and vary between 3.9% and 8.7% (excluding the systematic errors).

In Figure 2, we scrutinise more closely those samples relevant to the title of the comment by Guérin et al. (2013) – that is, samples associated with the HP and the SB. The beta dose rate adjustment was not applied to all samples; it was used to determine the OSL ages for eight of the 13 samples from Diepkloof (shown as crosses in Fig. 2; Jacobs et al., 2008a), seven of eight additional samples from Sibudu (shown as filled circles in Fig. 2; Jacobs et al., 2008a), and three of 10 samples from Blombos (shown as open triangles in Fig. 2; Jacobs et al., 2013). The open circles in Figure 2 are for the same Sibudu post-HP, late and final MSA ages shown in Fig. 1a,b. Figure 2 plots the difference in log ages against the mean log age for each sample (n = 31); scales of ages and ratios are also shown. We calculated an overall mean difference in log ages for all 31 samples

of 0.0233, with a standard deviation of 0.0412. This mean difference converts to an aFMM/CAM age ratio of 1.024 (95% confidence interval from 1.008 to 1.039), so the aFMM ages are, on average, ~2.4% larger than the CAM ages. The standard deviation of 0.0412 is an estimate of the standard error of an individual difference and the grey band in Figure 2 is calculated as 2.04 × 0.0412 (= 0.084) around zero, where 2.04 is the 97.5 percentage point of the t-distribution with 30 degrees of freedom (corresponding to 1.96 for a normal distribution). A point falling within this band is individually consistent (within estimation error) with zero difference (age ratio = 1). The band captures all of the points, except SIB9. The degree of concordance between each pair of log age estimates is unsurprising, as the aFMM ages are based on the D_e of the major component, which contains >83% of the grains for all but two (DRS6 and DRS9) of the 31 samples.

Guérin et al. (2013) suggest that the adjustment procedure of Jacobs et al. (2008b) could be improved by including an 'optimisation' feature that forces the ages of the two FMM components to be equivalent. This approach is sound in principle, but there are statistical arguments for and against the use of optimisation models (Galbraith, 2015b). If implemented, optimisation would need to account for the differing sizes of the standard errors of the two D_e components (often the minor component has a relative error several-fold larger than that of the major component – see, for example, Jacobs et al., 2012, their Tables 1 and 2) and the errors in the dose rate that are common to both components. Guérin et al. (2013) iterated their calculations until the age estimates for the two components were indistinguishable, but they do not say if they took the respective errors into account.

The effect of optimising the dose rate model is, anyway, very small: for the samples from El Harhoura 2 and El Mnasra, Guérin et al. (2013) obtained FMM ages that were, on average, ~1.5% lower after optimisation, which would draw them even closer to the CAM ages. We obtain a very similar result by adjusting the assumed beta dose rate of the minor D_e component until the mean ages of both FMM components match exactly. The mean of the log ratios is 0.013, corresponding to a ratio of 1.013 (95% confidence interval from 1.010 to 1.016), so the pre-optimisation ages are, on

average, higher than the post-optimisation ages by ~1.3%. The individual log ratios are all positive and vary between 0.005 and 0.031 (equivalent to ratios of between 1.005 and 1.031), with a standard deviation of 0.006. Again, these differences are very small compared with a typical relative standard error of a single aFMM age estimate of between 4.7% and 8.7% (excluding the systematic errors). It is also worth noting that not optimising the dose rate model is not necessarily invalid – that is, you can still obtain a valid estimate from the dominant component alone (Galbraith, 2015b).

The same conclusions hold true for other samples in which the major D_e component contains the vast majority of grains. Galbraith (2015b), however, elaborates on the statistical issues relating to dose rate heterogeneity, taking issue with the beta dose rate adjustment method of Jacobs et al. (2008b) and with several of the claims made on p. 315 of Guérin et al. (2013). More mathematically efficient and statistically rigorous approaches should, therefore, be used to refine the beta dose rate adjustment method of Jacobs et al. (2008b), and we welcome further efforts to this end. We acknowledge that our beta dose rate adjustment method has several shortcomings, so until direct measurements of beta dose rate at the single-grain scale, combined with improved models, are available to deal with issues of dose rate heterogeneity, our preferred approach at present is to use the CAM estimate of D_e and the average beta dose rate, rather than the aFMM estimate of D_e and the adjusted beta dose rate. What ultimately matters to archaeologists is to what extent the ages depend significantly on the choice of model – and we have demonstrated in this study that the answer is clearly ‘very little’ for these samples.

It is important to note that Figure 2 does not include two-thirds of the OSL ages presented in Jacobs et al. (2008a, 2013), because no beta dose rate adjustment was applied to 36 of the 54 samples – CAM or FMM ages were reported instead. The latter procedure was applied when two D_e components were identified with the FMM, but the minor dose component contained D_e values that were too small to be explained by beta dose heterogeneity (i.e., they represent younger intrusive grains), consisted of D_e values larger than the weighted mean D_e value of the major component

(e.g., DRS10), or where our field observations did not support an interpretation consistent with beta microdosimetry as the main cause of spread in D_e values (e.g., KKH samples). Based on the underlying assumptions of the adjustment method of Jacobs et al. (2008b), the beta dose rate correction can only be applied to a minor component composed of D_e values smaller than the main component. We interpret minor components with larger D_e values as representing intrusive older grains, but acknowledge that they may instead, or as well, represent a small proportion of grains exposed to higher beta dose rates than the sample average. Guérin et al. (2013: 316) err, therefore, in suggesting that Jacobs et al. (2008a) applied the beta dose rate adjustment model to Diepkloof sample DRS10, "resulting in an age significantly smaller than would otherwise be obtained." The adjustment method was not applied to this sample. Guérin et al. (2013: 315–316 and Fig. 1 caption) also presume that we incorrectly applied this method to Sibudu samples SIB2 and SIB7, by calculating the ages from the D_e values for the minor component; in fact, the ages of both samples were calculated using the procedure described by Jacobs et al. (2008b).

5.2 Revisiting the gap

Returning to the question of a time gap between the end of the SB and the start of the HP, Jacobs et al. (2008a: 19 in supplementary material) explicitly noted the possibility "that there was activity in this period and that samples from this time were simply not obtained". So, have any further studies in the intervening decade started to close or fill this gap? Work by one of us (Z.J.) at Pinnacle Point site 5-6 (PP5-6) on the south coast of South Africa shows clear evidence for 'filling the gap'. PP5-6 is a site that has a long, high-resolution depositional sequence that records major changes in stone artefact technology that have been argued to represent significant shifts in behaviour (Brown et al., 2009, 2012) – including features described by some as akin to Early and Intermediate HP at Diepkloof (e.g., Porraz et al., 2013; Conard and Porraz, 2015). Porraz et al. (2013) informally relate the SGS, OBS1 and SADBS stratigraphic aggregates at PP5-6 to, respectively, the Intermediate HP, MSA Jack and Early HP at Diepkloof. Site PP5-6 is exceptional

because it has sufficient resolution to allow natural sedimentary processes to be related directly to behavioural and cultural changes through a critical period of the MSA, aiding the unpacking of finer details of what constitutes and drives the different anthropogenic markers (Karkanas et al., 2015). The different types of sedimentary deposit preserved at PP5-6 also provide an ideal test of how well OSL dating can resolve the timing between clearly different stratigraphic aggregates, sub-aggregates and cultural phases.

Eleven ages have been published previously for this part of the PP5-6 sequence, which contains the HP and earlier microlithic industries of the SADBS (Brown et al., 2009, 2012). We have updated those ages here, so that they are consistent with the other ages presented in this study (Table 1). The updated ages are shown in Figure 3, together with the end date of the SB and start date of the HP (stippled lines) and the gap between them based on Estimate B (grey band). The different lithic phases follow a clear chronological succession, ending with the 'classic' HP in DBCS and OBS2, which has ages consistent with those for the HP in this study. The sequence includes the earlier phases of microlithic technology, including the SADBS with ages of between 66 ± 3 and 73 ± 3 ka (at 1σ), thus filling the gap between the SB and the HP. In the absence of the SB at PP5-6, the SADBS microlithic assemblage overlaps in age with the SB at the nearby site of Blombos, which can also be reconciled with the cultural sequence and overlap of bifacial points and HP backed artefacts at Diepkloof (Porraz et al., 2013). Filling the gap in this way reflects the greater resolution provided by site PP5-6, in particular, and by the systematic application of single-grain OSL dating to improve our ability to compare and integrate different data sets; it is not due to a change in methodology or our analytical or modelling procedures. Eliminating the gap between the SB and HP obviously has implications for how we define the HP and its duration.

We expect that fine-grained OSL chronologies can be obtained for different cultural phases at sites that are well-resolved stratigraphically. Taphonomic changes, however, commonly result in compression of the stratigraphy and time-averaging of anthropogenic contents, so an average depositional age for the sediments is the best that can be achieved using OSL dating. This restricts

the level of detail that can be discerned at sites that are less well preserved and resolved in space and time. Accordingly, our emphasis is now on dating sites such as PP5-6 at finer stratigraphic resolution, so that we can utilise the strengths of statistical models, such as the Bayesian models used widely in ^{14}C dating (e.g., Bronk Ramsey, 2009), to improve the precision of chronological sequences and interrogate cultural processes in more detail. Sites such as PP5-6 and Blombos will allow the HP and SB to be examined at higher temporal resolution, while other sites (e.g., Diepkloof and Sibudu) can provide insights into the chronological relationships of these and other lithic industries – and any spatial patterns – across southern Africa.

6. Conclusions

The modelled start and end ages for the SB and HP, the duration of the HP, and the length of the 'gap' between the SB and the HP all have confidence intervals that span several millennia. Owing to the number of separate measurements made to obtain estimates of the D_e and dose rate to calculate an OSL age, and the other uncertainties inherent to the method, it is not practicable to reduce the total error associated with individual or modelled ages to less than ~3%. At 60 ka, this represents 1800 years at 1σ or about 62 human generations, assuming that each generation is ~29 years (Fenner, 2005). This is much coarser than the decadal or centennial level resolution (corresponding to 1–3 generations) required to truly understand the operation of cultural processes, such as information flow, the impact of demographic changes on resident populations, or the effects of short-term environmental variations on resource use. It is fundamentally important, therefore, to be aware of these different scales of analysis when interpreting all types of data, including technological and stratigraphic changes through time.

The collection of more detailed data at higher stratigraphic resolution (e.g., Goldberg et al., 2009; Miller et al., 2013; Karkanas et al., 2015) and more comprehensive technological analyses are leading to changes in our understanding of the meaning of the HP and SB, and of the MSA more generally. This is evident from the finer details of diachronic changes within single assemblages

that have emerged from analyses of the HP at, for example, Klasies River (Wurz, 2002; Villa et al., 2010), Rose Cottage Cave (Soriano et al., 2007), PP5-6 (Brown et al., 2009, 2012); Klein Kliphuis (Mackay, 2010, 2011a), Diepkloof (Mackay, 2010, 2011b; Porraz et al., 2013), Klipdrift Rockshelter (Henshilwood et al., 2014) and Sibudu (Soriano et al., 2015). The same applies also to analyses of the SB (Archer et al., 2016) at Hollow Rock Shelter (Högberg and Larsson, 2011), Diepkloof (Porraz et al., 2013), Blombos (Archer et al., 2015; Soriano et al., 2015) and Sibudu (Soriano et al., 2015). These studies are improving the resolution of the questions we can ask of the HP and SB, and about the meaning of these labels to describe assemblages. Of increasing importance is the concept of scales of analysis and the implications of a more finely resolved understanding of the effects of taphonomic processes on our interpretation of chronological and other data used to reconstruct the human past.

Acknowledgements

We gratefully acknowledge the Australian Research Council for project support and fellowships to Z.J. (DP1092843) and R.G.R. (FL130100116). We especially thank Rex Galbraith for calculating the estimates in Tables 2 and 3, and Rex and Jane Galbraith, Alex Mackay, Lyn Wadley and Ann Wintle for constructive comments on an earlier version of this article. We also appreciate the feedback from the two anonymous reviewers, the Associate Editor and Guillaume Guérin.

References

Ankjærgaard, C., Murray, A.S., 2007. Total beta and gamma dose rates in trapped charge dating based on beta counting. Radiat. Meas. 42, 352–359.

Archer, W., Gunz, P., Van Niekerk, K.L., Henshilwood, C.S., McPherron, S.P., 2015. Diachronic change within the Still Bay at Blombos Cave, South Africa. PLOS ONE 10, e0132428,

Archer, W., Pop, C.M., Gunz, P., McPherron, S.P., 2016. What is Still Bay? Human biogeography and bifacial point variability. J. Hum. Evol. 97, 58–72.

Arnold, L.J., Roberts, R.G., 2009. Stochastic modelling of multi-grain equivalent dose (D_e) distributions: implications for OSL dating of sediment mixtures. Quatern. Geochron. 4, 204–230.

Box, G.E.P., 1979. Robustness in the strategy of scientific model building. In: R.L. Launer, G.N. Wilkinson (Eds), Robustness in Statistics. Academic Press, New York, pp. 201–236.

Bronk Ramsey, C., 2009. Bayesian analysis of radiocarbon dates. Radiocarbon 51, 337–360.

Brown, K.S., Marean, C.W., Herries, A.I.R., Jacobs, Z., Tribolo, C., Braun, D., Roberts, D.L., Meyer, M.C., Bernatchez, J., 2009. Fire as an engineering tool of early modern humans. Science 325, 859–862.

Brown, K.S., Marean, M.W., Jacobs, Z., Schoville, B.J., Oestmo, S., Fisher, E.C., Bernatchez, J., Karkanas, P., Matthews, T. 2012. An early and enduring advanced technology originating 71,000 years ago in South Africa. Nature 491, 590–594.

Conard, N.J., Porraz, G., 2015. Revising models for the cultural stratigraphic sequence of the Middle Stone Age. S. Afr. Archaeol. Bull.70, 127–130.

Cunningham, A.C., DeVries, D.J., Schaart, D.R., 2012. Experimental and computational simulation of beta-dose heterogeneity in sediment. Radiat. Meas. 47, 1060–1067.

Duller, G.A.T., 2008. Luminescence Dating: Guidelines on Using Luminescence Dating in Archaeology. English Heritage, Swindon.

Feathers, J., 2015. Luminescence dating at Diepkloof Rock Shelter – new dates from single-grain quartz. J. Archaeol Sci. 63, 164–174.

Fenner, J.N., 2005. Cross-cultural estimation of the human generation interval for use in genetics-based population divergence studies. Am. J. Phys. Anthropol. 128, 415–423.

Galbraith, R.F., 1988. Graphical display of estimates having differing standard errors. Technometrics 30, 271–281.

Galbraith, R.F., 2003. A simple homogeneity test for estimates of dose obtained using OSL. Ancient TL 21, 75–78.

Galbraith, R.F., 2005. Statistics for Fission Track Analysis. Chapman & Hall/CRC Press, Boca Raton.

Galbraith, R.F., 2015a. On the mis-use of mathematics: a comment on "How confident are we about the chronology of the transition between Howieson's Poort and Still Bay?" by Guérin et al. (2013). J. Hum. Evol. 80, 184–186.

Galbraith, R., 2015b. A note on OSL age estimates in the presence of beta dose heterogeneity. Ancient TL 33, 31–34.

Galbraith, R.F., Roberts, R.G., 2012. Statistical aspects of equivalent dose and error calculation and display in OSL dating: an overview and some recommendations. Quatern. Geochron. 11, 1–27.

Galbraith, R.F., Roberts, R.G., Laslett, G.M., Yoshida, H., Olley, J. M., 1999. Optical dating of single and multiple grains of quartz from Jinmium rock shelter, northern Australia: Part I, experimental design and statistical models. Archaeometry 41, 339–364.

Galbraith, R.F., Roberts, R.G., Yoshida, H., 2005. Error variation in OSL palaeodose estimates from single aliquots of quartz: a factorial experiment. Radiat. Meas. 39, 289–307.

Gliganic, L.A., Jacobs, Z., Roberts, R.G., Domínguez-Rodrigo, M., Mabulla, A.Z.P., 2012a. New ages for Middle and Later Stone Age deposits at Mumba rockshelter, Tanzania: optically stimulated luminescence dating of quartz and feldspar grains. J. Hum. Evol. 62, 533–547.

Gliganic, L.A., Jacobs, Z., Roberts, R.G., 2012b. Luminescence characteristics and dose distributions for quartz and feldspar grains from Mumba rockshelter, Tanzania. Archaeol. Anthropol. Sci. 4, 115–135.

Goldberg, P., Miller, C.E., Schiegl, S., Ligouis, B., Berna, F., Conard, N.J., Wadley, L., 2009. Bedding, hearths, and site maintenance in the Middle Stone Age of Sibudu Cave, KwaZulu-Natal, South Africa. Archaeol. Anthropol. Sci. 1, 95–122.

Grün, R., Beaumont, P.B., 2001. Border Cave revisited: a revised ESR chronology. J. Hum. Evol. 40, 467–482.

Guérin, G., Mercier, N., Adamiec, G., 2011. Dose-rate conversion factors: update. Ancient TL 29, 5–9.

Guérin, G., Murray, A.S., Jain, M., Thomsen, K.J., Mercier, N., 2013. How confident are we in the chronology of the transition between the Howieson's Poort and the Still Bay? J. Hum. Evol. 64, 314–317.

Henshilwood, C.S., d'Errico, F., Yates, R., Jacobs, Z., Tribolo, C., Duller, G.A.T., Mercier, N., Sealy, J.C., Valladas, H., Watts, I., Wintle, A.G., 2002. Emergence of modern human behaviour: Middle Stone Age engravings from South Africa. Science 295, 1278–1280.

Henshilwood, C.S., d'Errico, F., Vanhaeren, M., Van Niekerk, K., Jacobs, Z., 2004. Middle Stone Age shell beads from South Africa. Science 304, 404.

Henshilwood , C.S., Van Niekerk, K.L., Wurz, S., Delagnes, A., Armitage, S.J., Rifken, R.F., Douze, K., Keene, P., Haaland, M., Reynard, J., Discamps, E., Mienies, S.S., 2014. Klipdrif Shelter, southern Cape, South Africa: preliminary report on the Howiesons Poort layers. J. Archaeol Sci. 45, 284–303.

Högberg, A., Larsson, L., 2011. Lithic technology and behavioural modernity: new results from the Still Bay site, Hollow Rock Shelter, Western Cape Province, South Africa. J. Hum. Evol. 61, 133–155.

Jacobs, Z., Roberts, R.G., 2007. Advances in optically stimulated luminescence dating of individual grains of quartz from archeological deposits. Evol. Anthropol. 16, 210–223.

Jacobs, Z., Roberts, R.G., 2008. Testing times: old and new chronologies for the Howieson's Poort and Still Bay industries in environmental context. South African Archaeological Society Goodwin Series 10, 9–34.

Jacobs, Z., Roberts, R.G., 2009. Human history written in stone and blood. Am. Sci. 97, 302–309.

Jacobs, Z., Roberts, R.G., 2015. An improved single grain OSL chronology for the sedimentary deposits from Diepkloof Rockshelter, Western Cape, South Africa. J. Archaeol. Sci. 63, 175–192.

Jacobs, Z., Roberts, R.G., Galbraith, R.F., Deacon, H.J., Grün, R., Mackay, A., Mitchell, P., Vogelsang, R., Wadley, L., 2008a. Ages for the Middle Stone Age of southern Africa: implications for human behavior and dispersal. Science 322, 733–735.

Jacobs, Z., Wintle, A.G., Roberts, R.G., Duller, G.A.T., 2008b. Equivalent dose distributions from single grains of quartz at Sibudu, South Africa: context, causes and consequences for optical dating of archaeological deposits. J. Archaeol. Sci. 35, 1808–1820.

Jacobs, Z., Wintle, A.G., Duller, G.A.T., Roberts, R.G., Wadley, L., 2008c. New ages for the post-Howiesons Poort, late and final Middle Stone Age at Sibudu, South Africa. J. Archaeol. Sci. 35, 1790–1807.

Jacobs, Z., Meyer, M.C., Roberts, R.G., Aldeias, V., Dibble, H., El Hajraoui, M.A., 2011. Single-grain OSL dating at La Grotte des Contrebandiers ('Smugglers' Cave'), Morocco: improved age constraints for the Middle Paleolithic levels. J. Archaeol. Sci. 38, 3631–3643.

Jacobs, Z., Roberts, R.G., Nespoulet, R., El Hajraoui, M.A., Debénath, A., 2012. Single-grain OSL chronologies for Middle Palaeolithic deposits at El Mnasra and El Harhoura 2, Morocco: implications for Late Pleistocene human–environment interactions along the Atlantic coast of northwest Africa. J. Hum. Evol. 62, 377–394.

Jacobs, Z., Hayes, E.H., Roberts, R.G., Galbraith, R.F., Henshilwood, C.S., 2013. An improved OSL chronology for the Still Bay layers at Blombos Cave, South Africa: further tests of single-

grain dating procedures and a re-evaluation of the timing of the Still Bay industry across southern Africa. J. Archaeol. Sci. 40, 579–594.

Jankowski, N.R., Jacobs, Z., Goldberg, P., 2015. Optical dating and soil micromorphology at MacCauley's Beach, New South Wales, Australia. Earth Surface Processes and Landforms 40, 229–242.

Karkanas, P., Brown, K.S., Fisher, E.C., Jacobs, Z., Marean, C.W., 2015. Interpreting human behavior from depositional rates and combustion features through the study of sedimentary microfacies at site Pinnacle Point 5-6, South Africa. J. Hum. Evol. 85, 1–21.

Mackay, A., 2010. The late Pleistocene archaeology of Klein Kliphuis rock shelter, Western Cape, South Africa: 2006 excavations. S. Afr. Archaeol. Bull. 65, 132–147.

Mackay, A., 2011a. Nature and significance of the Howiesons Poort to post-Howiesons Poort transition at Klein Kliphuis rockshelter, South Africa. J. Archaeol. Sci. 38, 1430–1440.

Mackay, A., 2011b. Potentially stylistic differences between backed artefacts from two nearby sites occupied ~60,000 years before present in South Africa. J. Anthropol. Archaeol. 30, 235–245.

Mackay, A., Stewart, B.A., Chase, B.M., 2014. Coalescence and fragmentation in the late Pleistocene archaeology of southernmost Africa. J. Hum. Evol. 72, 26–51.

Martin, L., Incerti, S., Mercier, N., 2015a. DosiVox: implementing Geant 4-based software for dosimetry simulations relevant to luminescence and ESR dating techniques. Ancient TL 33, 1–10.

Martin, L., Mercier, N., Incerti, S., Lefrais, Y., Pecheyran, C., Guérin, G., Jarry, M., Bruxelles, L., Bon, F., Pallier, C., 2015b. Dosimetric study of sediments at the beta dose rate scale: characterization and modelization with the DosiVox software. Radiat. Meas. 81, 134–141.

Miller, C.E., Goldberg, P., Berna, F., 2013. Geoarchaeological investigations at Diepkloof Rock Shelter, Western Cape, South Africa. J. Archaeol. Sci. 40, 3432–3452.

Miller, G.H., Beaumont, P.B., Deacon, H.J., Brooks, A.S., Hare, P.E., Jull, A.J.T., 1999. Earliest modern humans in South Africa dated by isoleucine epimerization in ostrich eggshell. Quaternary Sci. Rev. 18, 1537–1548.

Murray, A.S., Olley, J.M., 2002. Precision and accuracy in the optically stimulated luminescence dating of sedimentary quartz: a status review. Geochronometria 21, 1–16.

Murray-Wallace, C.V., Richter, J., Vogelsang, R., 2015. Aminostratigraphy and taphonomy of ostrich eggshell in the sedimentary infill of Apollo 11 Rockshelter, Namibia. J. Archaeol. Sci.: Reports 4, 143–151.

Nathan, R.P., Thomas, P.J., Jain, M., Murray, A.S., Rhodes, E.J., 2003. Environmental dose rate heterogeneity of beta radiation and its implications for luminescence dating: Monte Carlo modelling and experimental validation. Radiat. Meas. 37, 305–313.

Oldham, P.D., 1962. A note on the analysis of repeated measurements of the same subjects. J. Chron. Dis. 15, 969–977.

Olley, J.M., Pietsch, T., Roberts, R.G., 2004. Optical dating of Holocene sediments from a variety of geomorphic settings using single grains of quartz. Geomorphology 60, 337–358.

Porraz, G., Texier, P-J., Archer, W., Piboule, M., Rigaud, J.-P., Tribolo, C., 2013. Technological successions in the Middle Stone Age sequence of Diepkloof Rock Shelter, Western Cape, South Africa. J. Archaeol. Sci. 40, 3376–3400.

Roberts, R.G., Galbraith, R.F., Yoshida, H., Laslett, G.M., Olley, J.M., 2000. Distinguishing dose populations in sediment mixtures: a test of single-grain optical dating procedures using mixtures of laboratory-dosed quartz. Radiat. Meas. 32, 459–465.

Roberts, R.G., Jacobs, Z., Li, B., Jankowski, N.R., Cunningham, A.C., Rosenfeld, A.B., 2015. Optical dating in archaeology: thirty years in retrospect and grand challenges for the future. J. Archaeol. Sci. 56, 41–60.

Schiegl, S., Conard, N., 2006. The Middle Stone Age sediments at Sibudu: results from FTIR spectroscopy and microscopic analyses. Southern African Humanities 18, 149–172.

Schiegl, S., Stockhammer, P., Scott, C., Wadley, L., 2004. A mineralogical and phytolith study of the Middle Stone Age hearths in Sibudu Cave, KwaZulu-Natal, South Africa. S. Afr. J. Sci. 100, 185–194.

Soriano, S., Villa, P., Wadley, L., 2007. Blade technology and tool forms in the Middle Stone Age of South Africa: the Howiesons Poort and post-Howiesons Poort at Rose Cottage Cave. J. Archaeol. Sci. 34, 681–703.

Soriano, S., Villa, P., Delagnes, A., Degano, I., Pollarolo, L., Lucejko, J.J., Henshilwood, C., Wadley, L., 2015. The Still Bay and Howiesons Poort at Sibudu and Blombos: understanding Middle Stone Age technologies. PLOS ONE 10, e0131127.

Texier, P.-J., Porraz, G., Parkington, J., Rigaud, J.-P., Poggenpoel, C., Miller, C., Tribolo, C., Cartwright, C., Coudenneau, A., Klein, R., Steele, T., Verna, C., 2010. A Howiesons Poort tradition of engraving ostrich eggshell containers dated to 60,000 years ago at Diepkloof Rock Shelter, South Africa. Proc. Natl. Acad. Sci. 107, 6180–6185.

Texier, P.-J., Porraz, G., Parkington, J., Rigaud, J.-P., Poggenpoel, C., Tribolo, C., 2013. The context, form and significance of the MSA engraved ostrich eggshell collection from Diepkloof Rock Shelter, Western Cape, South Africa. J. Archaeol. Sci. 40, 3412–3431.

Tribolo, C., Mercier, N., Valladas, H., 2005. Chronology of the Howieson's Poort and Still Bay technocomplexes: assessment and new data from luminescence. In: D'Errico, F., Backwell, L. (Eds), From Tools to Symbols: From Early Hominids to Modern Humans. Witwatersrand University Press, Johannesburg, pp. 493–511.

Tribolo, C., Mercier, N., Selo, M., Valladas, H., Joron, J.-L., Reyss, J.-L., Henshilwood, C.S., Sealy, J.C., Yates, R., 2006. TL dating of burnt lithics from Blombos Cave (South Africa): further evidence for the antiquity of modern human behaviour. Archaeometry 48, 341–357.

Tribolo, C., Mercier, N., Douville, E., Joron, J.-L., Reyss, J.-L., Rufer, D., Cantin, N., Lefrais, Y., Miller, C.E., Parkington, J., Porraz, G., Rigaud, J.-P., Texier, P.-J., 2013. OSL and TL dating of the Middle Stone Age sequence of Diepkloof Rock Shelter (Western Cape, South Africa): a clarification. J. Archaeol. Sci. 40, 3401–3411.

Valladas, H., Wadley, L., Mercier, N., Tribolo, C., Reyss, J.L., Joron, J.L., 2005. Thermoluminescence dating on burnt lithics from Middle Stone Age layers at Rose Cottage Cave. S. Afr. J. Sci. 101, 169–174.

Vanhaeren, M., D'Errico, F., van Niekerk, K.L., Henshilwood, C.S., Erasmus, R.M., 2013. Thinking strings: additional evidence for personal ornament use in the Middle Stone Age at Blombos Cave, South Africa. J. Hum. Evol. 64, 500–517.

Villa, P., Soriano, S., Teyssandier, N., Wurz, S., 2010. The Howiesons Poort and MSA III at Klasies River main site, Cave 1A. J. Archaeol. Sci. 37, 630–655.

Vogel, J.C., 2001. Radiometric dates for the Middle Stone Age in South Africa. In: P.V. Tobias, M.A. Raath, J. Moggi-Cecchi, G.A. Doyle (Eds), Humanity from African Naissance to Coming Millennia: Colloquia in Human Biology and Palaeoanthropology. Firenze University Press, Witwatersrand University Press, Johannesburg, pp. 261–268.

Wadley, L., Hodgskiss, T., Grant, M., 2009. Implications for complex cognition from the hafting of tools with compound adhesives in the Middle Stone Age, South Africa. Proc. Natl. Acad. Sci.106, 9590–9594.

Wadley, L., Sievers, C., Bamford, M., Goldberg, P., Berna, F., Miller, C., 2011. Middle Stone Age bedding construction and settlement patterns at Sibudu, South Africa. Science 334, 1388–1391.

Wintle, A.G., 2014. Luminescence dating methods. In: H.D. Holland, K.K. Turekian (Eds), Treatise on Geochemistry (2nd edition), vol. 14, Elsevier, Oxford, pp. 17–35.

Wurz, S., 2002. Variability in the Middle Stone Age lithic sequence, 115,000–60,000 years ago at Klasies River, South Africa. J. Archaeol. Sci. 29, 1001–1015.

Figure captions

Figure 1: a) Comparison of aFMM and CAM (see definitions in text) ages calculated for the same samples from Sibudu; the error bars are σ1 in the notation used here. SIB12 is not shown since only a CAM age was calculated for this sample. This figure shows the same data as Figure 1 of Guérin et al. (2013), but using the dose rate errors as published by Jacobs et al. (2008a), rather than the 7% applied to all beta dose rates by Guérin et al. (2013; see further discussion in our main text). b) Corresponding plot of the difference between each pair of log age estimates (on the vertical axis) against the mean log age estimate of the pair (on the horizontal axis). The grey band is calculated as 2.18 estimated standard deviations above and below zero (based on 12 degrees of freedom). So, for a point within this band, the 95% confidence interval for the true difference includes zero (see text).

Figure 2: Comparison of OSL age estimates obtained by two methods (aFMM and CAM; see definitions in text) for 31 samples from four groups. The difference in log age is plotted against the mean log age for each sample. Scales of ages and their ratios are also shown. The grey band denotes ± 2.04 standard deviations about zero for individual differences (with 30 degrees of freedom), so each point within this band has a 95% confidence interval for the true difference that includes zero.

Figure 3: Published ages for the Pinnacle Point site 5-6 (PP5-6) deposits that contain 'classic' HP (DBCS) and other microlithic industries (OBS1 and SADBS). The error bars (σ2) denote the 95% confidence interval for each age. The stippled lines are the end and start ages of the Still Bay (SB) and Howieson's Poort (HP), respectively, and the grey band marks the gap between the SB and HP based on Estimate B in Tables 2 and 3.

Table

Table 1: Dose rate data, D_e values and OSL ages (in ka) for all samples in this study[a].

Sample	Water content (%)	Environmental dose rate (Gy/ka)				D_e value (Gy)	Over-dispersion (%)	Age model [c]	Number of grains	Age ± σ2 (σ1) [d]
		Beta [b]	Gamma	Cosmic	Total					
Apollo 11										
AP-9	3 ± 1	1.97 ± 0.09	0.91 ± 0.04	0.04	2.95 ± 0.10	165.0 ± 7.0	5 ± 8	CAM	49/500	55.9 ± 3.0 (2.4)
AP-4	3 ± 1	1.60 ± 0.07	0.91 ± 0.04	0.04	2.59 ± 0.08	167.6 ± 4.8	17 ± 2	CAM	201/1000	64.7 ± 2.8 (1.9)
AP-5	3 ± 1	1.93 ± 0.08	0.92 ± 0.04	0.04	2.92 ± 0.09	193.7 ± 4.8	6 ± 4	CAM	64/500	66.3 ± 3.0 (2.1)
AP-6	3 ± 1	1.63 ± 0.07	1.12 ± 0.04	0.04	2.83 ± 0.09	197.7 ± 3.3	16 ± 2	CAM	15/1000	70.0 ± 2.9 (1.9)
Blombos Cave										
BBC10-1	6 ± 2	0.57 ± 0.03	0.42 ± 0.01	0.04	1.06 ± 0.04	78.8 ± 2.7	15 ± 3	CAM	111/1000	74.1 ± 3.6 (2.9)
BBC10-2	6 ± 2	0.58 ± 0.03	0.47 ± 0.02	0.04	1.12 ± 0.04	81.6 ± 2.1	21 ± 2	CAM	391/1000	72.8 ± 3.1 (2.3)
ZB4	11 ± 3	0.57 ± 0.03	0.44 ± 0.02	0.04	1.08 ± 0.04	77.6 ± 2.7	18 ± 3	CAM	122/1000	71.9 ± 3.7 (3.1)
BBC10-3	7 ± 2	0.61 ± 0.03	0.44 ± 0.01	0.04	1.12 ± 0.04	83.1 ± 3.0	20 ± 3	CAM	110/1000	74.3 ± 3.8 (3.1)
BBC10-4	6 ± 2	0.59 ± 0.03	0.38 ± 0.01	0.04	1.04 ± 0.04	74.6 ± 2.4	17 ± 3	CAM	136/1000	72.0 ± 3.4 (2.7)
BBC10-5	8 ± 2	0.54 ± 0.03	0.44 ± 0.01	0.04	1.05 ± 0.04	77.7 ± 2.3	26 ± 2	CAM	268/1000	74.1 ± 3.4 (2.6)
BBC10-6	13 ± 3	0.56 ± 0.03	0.47 ± 0.02	0.04	1.10 ± 0.04	76.0 ± 2.9	28 ± 3	CAM	119/1000	69.4 ± 3.8 (3.3)
ZB10	6 ± 2	0.58 ± 0.03	0.40 ± 0.03	0.04	1.04 ± 0.04	79.3 ± 2.5	23 ± 3	CAM	113/1000	76.0 ± 3.8 (3.2)
BBC10-7	20 ± 5	0.67 ± 0.05	0.48 ± 0.03	0.04	1.21 ± 0.05	87.7 ± 2.1	20 ± 3	CAM	136/1000	72.4 ± 3.7 (3.0)
BBC10-8	10 ± 3	0.65 ± 0.04	0.44 ± 0.03	0.04	1.16 ± 0.05	83.8 ± 2.4	26 ± 2	CAM	334/2000	72.3 ± 3.5 (2.8)
Diepkloof Rockshelter										
DRS9	5 ± 2	1.40 ± 0.07	0.70 ± 0.04	0.04	2.17 ± 0.08	94.6 ± 2.3	27 ± 1	CAM	569/1000	43.6 ± 1.9 (1.0)
DRS8	5 ± 2	1.17 ± 0.06	0.59 ± 0.03	0.04	1.83 ± 0.07	94.8 ± 2.4	25 ± 1	CAM	423/1000	51.9 ± 2.3 (1.4)
DRS6	5 ± 2	1.28 ± 0.06	0.55 ± 0.04	0.04	1.90 ± 0.07	112.5 ± 2.8 93.4 ± 2.4	34 ± 1	FMM CAM	538/1000	59.2 ± 2.7 (1.6) 49.1 ± 2.2 (1.3)
DRS5	5 ± 2	1.40 ± 0.07	0.57 ± 0.03	0.04	2.03 ± 0.07	113.7 ± 2.9	33 ± 2	FMM	425/1000	56.0 ± 2.5 (1.3)
DRS4	5 ± 2	1.16 ± 0.06	0.60 ± 0.03	0.04	1.82 ± 0.07	112.2 ± 2.6	21 ± 1	FMM	474/1000	61.6 ± 2.7 (1.5)
DRS7	5 ± 2	1.33 ± 0.06	0.75 ± 0.04	0.04	2.14 ± 0.07	123.1 ± 3.1	32 ± 2	FMM	375/1000	57.5 ± 2.6 (1.5)
DRS10	5 ± 2	0.98 ± 0.05	0.47 ± 0.03	0.04	1.52 ± 0.06	100.1 ± 2.4	23 ± 1	CAM	734/1000	65.9 ± 3.0 (1.7)
DRS11	5 ± 2	1.35 ± 0.06	0.63 ± 0.03	0.03	2.04 ± 0.07	124.2 ± 3.2	25 ± 1	CAM	461/1000	60.8 ± 2.6 (1.4)
DRS13	5 ± 2	1.10 ± 0.06	0.50 ± 0.02	0.03	1.66 ± 0.06	108.3 ± 2.5	26 ± 1	CAM	808/2000	65.1 ± 2.8 (1.5)
DRS14	5 ± 2	0.75 ± 0.04	0.47 ± 0.02	0.03	1.29 ± 0.05	98.5 ± 2.2	24 ± 1	CAM	1040/2400	76.5 ± 3.3 (2.4)

DRS15	5 ± 2	0.70 ± 0.04	0.41 ± 0.03	0.03	1.18 ± 0.05	104.4 ± 2.4	23 ± 1	CAM	791/2000	88.4 ± 4.0 (2.6)
DRS16	5 ± 2	0.57 ± 0.03	0.38 ± 0.02	0.03	1.01 ± 0.04	94.3 ± 2.7	23 ± 2	CAM	265/1000	93.3 ± 4.4 (3.3)
DRS17	5 ± 2	0.58 ± 0.03	0.51 ± 0.04	0.03	1.14 ± 0.05	100.8 ± 3.0	26 ± 2	CAM	255/1000	88.2 ± 4.4 (3.3)
Klein Kliphuis										
KKH8	3 ± 1	0.95 ± 0.05	0.60 ± 0.02	0.06	1.65 ± 0.05	53.7 ± 1.3	24 ± 2	FMM	258/1000	32.6 ± 1.3 (0.9)
KKH4	3 ± 1	0.81 ± 0.04	0.65 ± 0.02	0.05	1.54 ± 0.05	86.6 ± 2.7	32 ± 2	FMM	190/1000	56.1 ± 2.4 (1.8)
KKH3	3 ± 1	0.66 ± 0.04	0.56 ± 0.02	0.05	1.29 ± 0.04	78.2 ± 2.0	22 ± 2	FMM	261/1000	60.4 ± 2.5 (1.8)
KKH1	3 ± 1	0.85 ± 0.04	0.61 ± 0.02	0.04	1.54 ± 0.05	91.4 ± 2.6	39 ± 2	FMM	245/1000	59.4 ± 2.5 (1.8)
KKH5	3 ± 1	0.66 ± 0.04	0.56 ± 0.03	0.05	1.30 ± 0.05	86.4 ± 2.0	20 ± 2	FMM	279/1000	66.7 ± 3.0 (2.3)
KKH6	3 ± 1	1.00 ± 0.05	0.68 ± 0.03	0.04	1.76 ± 0.06	103.8 ± 2.6	23 ± 2	FMM	252/1000	59.1 ± 2.4 (1.7)
KKH7	3 ± 1	0.77 ± 0.04	0.62 ± 0.03	0.04	1.46 ± 0.05	91.2 ± 2.3	21 ± 2	FMM	213/1000	62.6 ± 2.6 (2.0)
KKH2	3 ± 1	0.70 ± 0.04	0.65 ± 0.02	0.04	1.42 ± 0.05	97.0 ± 2.5	21 ± 2	FMM	236/1000	68.2 ± 2.8 (2.0)
Klasies River										
ZKR5	8 ± 2	0.68 ± 0.03	0.49 ± 0.01	0.10	1.30 ± 0.05	76.8 ± 2.2	14 ± 2	CAM	145/1000	59.3 ± 2.8 (2.5)
ZKR9	15 ± 4	0.77 ± 0.05	0.69 ± 0.04	0.06	1.55 ± 0.06	98.2 ± 2.4	18 ± 2	FMM	256/2000	63.3 ± 2.9 (2.3)
ZKR10	8 ± 2	0.50 ± 0.03	0.36 ± 0.02	0.11	1.00 ± 0.04	66.8 ± 1.9	18 ± 2	FMM	151/1400	66.9 ± 3.3 (2.5)
ZKR6	15 ± 4	0.48 ± 0.02	0.35 ± 0.02	0.04	0.92 ± 0.03	58.1 ± 1.5	18 ± 2	FMM	220/1900	63.2 ± 2.7 (2.0)
ZKR4	15 ± 4	0.34 ± 0.04	0.32 ± 0.02	0.05	0.74 ± 0.04	53.4 ± 1.7	26 ± 3	FMM	113/1000	71.8 ± 4.7 (4.3)
ZKR8	15 ± 4	0.43 ± 0.05	0.35 ± 0.02	0.11	0.92 ± 0.08	64.0 ± 2.1	25 ± 3	FMM	126/1000	69.9 ± 6.2 (6.0)
Melikane										
MLK4	19 ± 5	1.18 ± 0.07	0.69 ± 0.04	0.08	1.98 ± 0.09	97.4 ± 2.7	14 ± 2	CAM	103/800	49.2 ± 2.5 (2.0)
MLK3	19 ± 5	1.25 ± 0.08	0.78 ± 0.04	0.08	2.15 ± 0.09	127.7 ± 3.7	9 ± 3	CAM	64/1000	59.5 ± 3.3 (2.8)
MLK2	19 ± 5	1.37 ± 0.09	0.87 ± 0.05	0.08	2.35 ± 0.10	183.7 ± 5.3	25 ± 2	FMM	251/1000	78.2 ± 4.0 (3.3)
Ntloana Tsoana										
NT2	20 ± 5	1.12 ± 0.08	0.54 ± 0.03	0.11	1.80 ± 0.08	99.8 ± 2.8	20 ± 2	FMM	167/1000	55.5 ± 3.0 (2.4)
NT1	10 ± 3	1.30 ± 0.07	0.59 ± 0.03	0.11	2.03 ± 0.07	121.4 ± 4.6	0	CAM	33/1000	59.8 ± 3.2 (2.5)
Rose Cottage Cave										
RCC5	12 ± 3	1.11 ± 0.06	0.52 ± 0.02	0.05	1.71 ± 0.07	94.3 ± 3.6	8 ± 5	CAM	50/900	55.1 ± 3.0 (2.4)
RCC2	12 ± 3	1.23 ± 0.07	0.57 ± 0.03	0.05	1.88 ± 0.07	120.7 ± 3.1	15 ± 2	FMM	194/1000	64.2 ± 3.0 (2.1)
RCC1	12 ± 3	1.25 ± 0.07	0.55 ± 0.05	0.05	1.88 ± 0.07	121.5 ± 3.2	22 ± 3	CAM	98/1000	64.6 ± 3.0 (2.2)
RCC3	12 ± 3	1.07 ± 0.06	0.57 ± 0.03	0.05	1.71 ± 0.07	111.4 ± 4.3	23 ± 3	CAM	92/1000	65.0 ± 3.5 (2.9)

Sehonghong										
SEH4	10 ± 3	1.02 ± 0.05	0.60 ± 0.03	0.06	1.72 ± 0.06	53.6 ± 1.8	29 ± 3	FMM	84/1000	31.2 ± 1.5 (1.2)
SEH3	6 ± 2	1.42 ± 0.07	0.77 ± 0.03	0.06	2.29 ± 0.07	68.5 ± 7.4	3 ± 3	CAM	12/1000	30.0 ± 3.4 (3.2)
SEH2	6 ± 2	1.44 ± 0.07	0.85 ± 0.03	0.06	2.39 ± 0,08	109.5 ± 5.3	15 ± 5	CAM	39/1000	45.9 ± 2.7 (2.3)
SEH1	10 ± 3	1.74 ± 0.09	0.81 ± 0.04	0.06	2.65 ± 0.10	129.8 ± 4.4	14 ± 3	CAM	64/1000	49.0 ± 2.4 (1.9)
Sibudu										
SIB9	5 ± 2	0.93 ± 0.05	0.77 ± 0.04	0.19	1.92 ± 0.07	96.5 ± 2.8	30 ± 2	CAM	141/	50.2 ± 2.5 (2.0)
SIB15	5 ± 2	1.18 ± 0.06	0.59 ± 0.03	0.17	1.97 ± 0.07	114.2 ± 3.0	24 ± 1	CAM	303/1100	58.0 ± 2.5 (1.7)
SIB17	5 ± 2	1.08 ± 0.05	0.66 ± 0.04	0.17	1.94 ± 0.07	117.4 ± 3.2	21 ± 2	CAM	196/1000	60.5 ± 2.8 (2.0)
SIB19	5 ± 2	1.28 ± 0.06	0.68 ± 0.04	0.15	2.13 ± 0.07	129.1 ± 3.4	32 ± 2	CAM	280/900	60.5 ± 2.6 (1.8)
SIB20	5 ± 2	1.08 ± 0.05	0.66 ± 0.04	0.14	1.91 ± 0.07	132.4 ± 3.4	23 ± 1	CAM	316/1000	69.5 ± 3.0 (2.1)
SIB21	5 ± 2	1.60 ± 0.08	0.73 ± 0.04	0.13	2.49 ± 0.09	176.5 ± 4.6	23 ± 1	CAM	272/1000	70.9 ± 3.2 (2.2)
SIB24[e]	5 ± 2	1.15 ± 0.06	0.73 ± 0.04	0.13	2.04 ± 0.11	144.5 ± 3.5	20 ± 1	CAM	315/1000	70.8 ± 4.1 (3.2)
SIB23[e]	5 ± 2	1.44 ± 0.07	0.82 ± 0.04	0.13	2.43 ± 0.08	176.0 ± 4.5	23 ± 1	CAM	302/1000	72.4 ± 3.0 (1.9)
Pinnacle Point site 5-6 (PP5-6)										
46791	8 ± 2	0.89 ± 0.04	0.57 ± 0.02	0.09	1.57 ± 0.05	93.2 ± 2.3	16 ± 3	CAM	127/900	59.3 ± 2.7 (2.1)
46790	5 ± 2	0.80 ± 0.04	0.55 ± 0.02	0.09	1.47 ± 0.05	90.4 ± 4.5	19 ± 5	CAM	36/900	61.5 ± 3.8 (3.4)
46789	5 ± 2	0.74 ± 0.04	0.58 ± 0.02	0.09	1.43 ± 0.05	91.4 ± 1.9	0	CAM	100/1000	63.7 ± 2.7 (2.0)
162513	4 ± 1	0.74 ± 0.03	0.57 ± 0.02	0.08	1.42 ± 0.04	96.8 ± 3.2	18 ± 4	CAM	101/2000	68.1 ± 3.3 (2.7)
46788	8 ± 2	0.54 ± 0.03	0.47 ± 0.02	0.08	1.12 ± 0.04	73.6 ± 1.5	15 ± 2	CAM	158/1000	65.9 ± 3.0 (2.3)
110633	4 ± 1	0.56 ± 0.03	0.47 ± 0.02	0.08	1.15 ± 0.04	78.3 ± 3.0	34 ± 4	CAM	89/1000	68.2 ± 3.7 (3.2)
46787	8 ± 2	0.62 ± 0.03	0.47 ± 0.02	0.08	1.20 ± 0.04	85.3 ± 1.8	19 ± 2	CAM	198/900	71.4 ± 3.2 (2.4)
46786	12 ± 3	0.51 ± 0.03	0.44 ± 0.02	0.08	1.06 ± 0.04	75.0 ± 2.9	21 ± 2	CAM	198/900	70.6 ± 4.0 (3.5)
162511	5 ± 2	0.69 ± 0.03	0.53 ± 0.02	0.08	1.32 ± 0.04	87.7 ± 3.3	23 ± 4	CAM	93/1000	66.6 ± 3.5 (3.0)
46785	5 ± 2	0.63 ± 0.03	0.55 ± 0.02	0.08	1.28 ± 0.04	91.2 ± 1.9	22 ± 2	CAM	246/900	71.1 ± 3.0 (2.2)
162510	5 ± 2	0.52 ± 0.03	0.53 ± 0.02	0.08	1.15 ± 0.04	84.2 ± 2.3	23 ± 3	CAM	174/1000	73.2 ± 3.4 (2.7)

[a]Ages listed here are the preferred ages included in Estimate B (except PP5-6; see text for details). CAM and FMM defined in text.

[b] Beta dose rates are all based on the sample-average beta dose rate obtained from GM-25-5 beta counting; see Jacobs and Roberts (2015) for details.

[c] The FMM was used only when samples showed evidence of mixing (see Jacobs et al., 2008a; Jacobs and Roberts, 2015).

[d] Uncertainties are given at 1σ (standard error of the mean). The σ2 values listed here were derived by combining, in quadrature, all known and estimated sources of random and systematic error, whereas the σ1 values exclude the systematic errors (see text).

[e] The dose rate data for these two samples were transposed in Jacobs et al. (2008a) and are shown correctly here.

Table 2: Maximum likelihood estimates and 95% confidence intervals for the start and end ages (in ka) of the Howieson's Poort (HP) and Still Bay (SB) periods[a].

	End of HP		Start of HP		End of SB		Start of SB	
	Age	**95% CI**	**Age**	**95% CI**	**Age**	**95% CI**	**Age**	**95% CI**
Original	59.5	56.5–62.7	64.8	61.6–68.2	71.3	67.3–74.69	71.9	68.3–75.7
Estimate A	58.7	55.9–62.5	66.5	62.0–70.9	70.6	65.1–76.0	73.2	68.6–78.7
Estimate B	58.3	54.5–62.0	66.6	62.2–71.1	71.9	66.0–77.8	72.2	67.2–77.2
Estimate C	58.2	54.4–61.9	66.8	62.2–71.3	71.6	66.0–77.1	71.9	67.0–76.7

[a]The 'Original' estimates are those published in Jacobs et al. (2008a). The three sets of 'Estimates' (A, B and C) are explained in the text. In addition to measurement and estimation errors, the confidence intervals for Estimates A, B and C include a relative systematic error of 3%, rather than 2% in the Original estimates. Note that the confidence intervals for different start and end ages are highly dependent and cannot be compared together.

Table 3: Maximum likelihood estimates and 95% confidence intervals for the durations (in ka) of the Howieson's Poort (HP) period and of the gap between the end of the Still Bay (SB) and start of the HP[a].

	Duration of HP period		**Duration of gap between SB and HP**		
	Estimate	**95% CI**	**Estimate**	**95% CI**	***p*-value**
Original	5.3	2.0–8.3	6.7	2.7–9.3	0.008
Estimate A	7.7	4.9–10.5	4.1	–0.1–8.3	0.059
Estimate B	8.4	5.6–11.1	5.3	0.7–9.9	0.014
Estimate C	8.6	5.7–11.6	4.8	0.5–9.1	0.033

[a]Because we are estimating differences between ages, systematic errors common to all estimates do not contribute to the uncertainties in these estimates. The final column gives *p*-values for testing the null hypothesis that there is no gap between the SB and HP; the smaller the *p*-value, the stronger is the evidence in favour of a gap.

Figure 1

Click here to download high resolution image

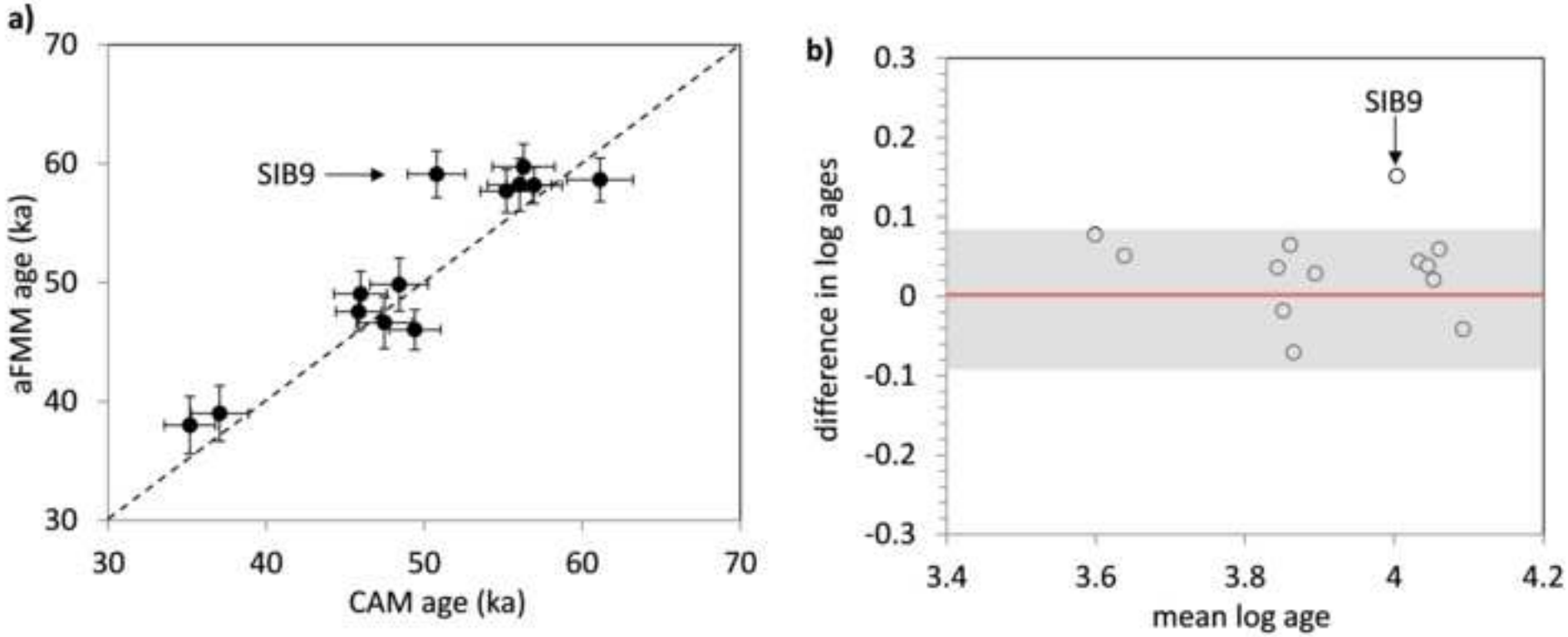


Figure 2

Click here to download high resolution image

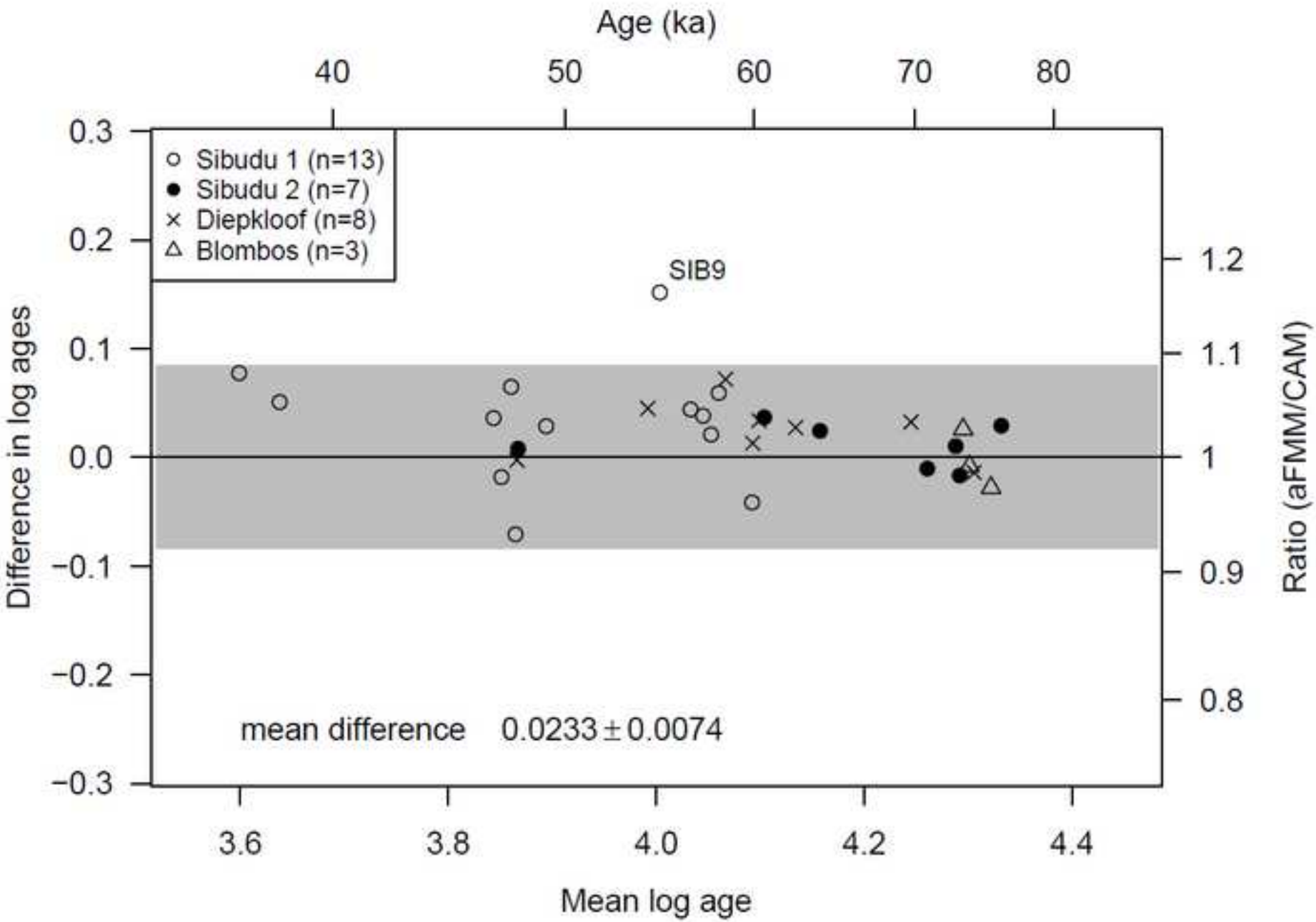
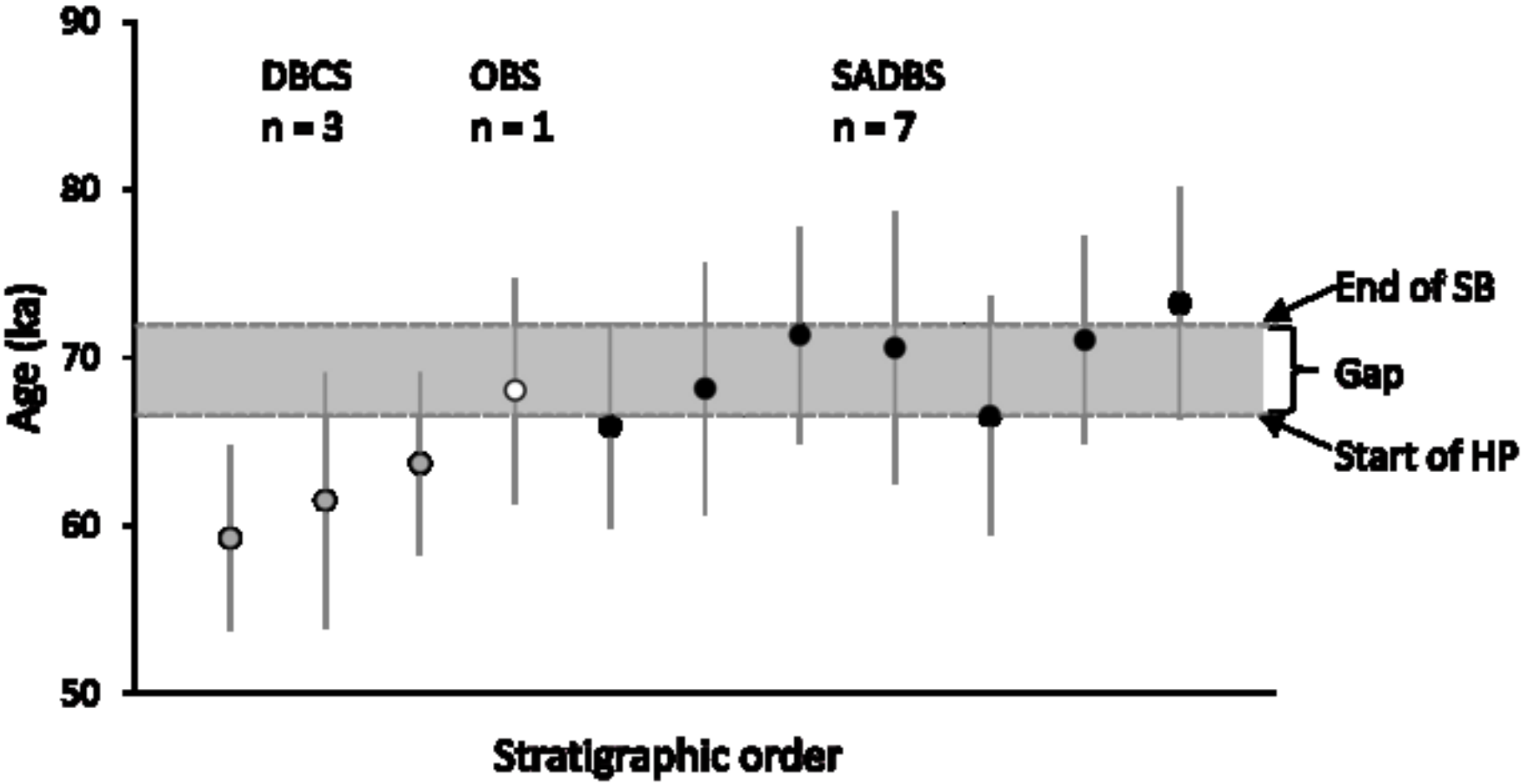


Figure 3

Click here to download high resolution image



Supplementary Material

Click here to download Supplementary Material: Supplementary Online Material Jacobs and Roberts.docx